

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:09:23 ; Search time 199 Seconds
(without alignments)
397.428 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRGTGSTGADGPGG.....WITQCFLPVFLAQPFGQR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_21:**
1: Geneseqp1980s:**
2: Geneseqp1990s:**
3: Geneseqp2000s:**
4: Geneseqp2001s:**
5: Geneseqp2002s:**
6: Geneseqp2003as:**
7: Geneseqp2003bs:**
8: Geneseqp2004s:**
9: Geneseqp2005s:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	959	100.0	180	2	AAW62584	Cancer as
2	959	100.0	180	2	AAW69665	Human NY-
3	959	100.0	180	2	AAW05965	Human can
4	959	100.0	180	3	AAW52430	Human tum
5	959	100.0	180	3	AAW70862	Human oes
6	959	100.0	180	3	AAW03154	Human NY-
7	959	100.0	180	4	AAW69946	Human NY-
8	959	100.0	180	4	AAW67164	Amino aci
9	959	100.0	180	4	AAW01535	Human NY-
10	959	100.0	180	4	AAW07714	Human NY-
11	959	100.0	180	5	AAW4818	Human NYN
12	959	100.0	180	5	AAW11543	Human tum
13	959	100.0	180	6	AAW58672	Human can
14	959	100.0	180	6	AAW48210	Human bla
15	959	100.0	180	6	AAW56508	Lung can
16	959	100.0	180	6	AAW56694	Lung can
17	959	100.0	180	6	AAW74198	Human NY-
18	959	100.0	180	6	AAW83438	Human NY-
19	959	100.0	180	7	AAW09576	NY-ESO-1
20	959	100.0	180	7	AAW09576	Human NY-
21	959	100.0	180	7	AAW09576	Human NY-
22	959	100.0	180	7	AAW09576	Binding d
23	959	100.0	180	8	AAW54139	Human NY-
24	959	100.0	180	8	AAW72815	Human NY-

25	959	100.0	180	8	ADQ18451	Human sof
26	959	100.0	180	8	ADQ10446	Autoimmun
27	959	100.0	180	8	ADW80926	Tumour as
28	959	100.0	180	9	ADW44353	Human auc
29	959	100.0	180	9	ADW85096	Tumour ant
30	959	100.0	180	9	AEA35651	Human NY-
31	959	100.0	397	4	AAE13122	NY-ESO-1/
32	954	99.5	240	9	ADW99402	NY-ESO-1/
33	954	99.5	240	9	AEW80047	Human NY-
34	953	98.4	180	6	AEU64816	Human NY-
35	953	99.4	180	9	ADZ28913	NY-ESO-1
36	953	99.4	180	9	ADZ42374	Immunogen
37	950	99.1	180	8	ADM73418	CAG-3 pro
38	950	99.1	180	8	ADM73417	Human NY-
39	934.5	97.4	179	8	ADK68648	Epitope 1
40	918	95.7	180	7	ADD35564	Human NY-
41	821	85.6	180	6	ABF74199	Human LAG
42	821	85.6	180	7	ADC09577	LAGR-1a p
43	821	85.6	180	8	ADM72816	Human LAG
44	814	84.9	180	2	AAW69664	Human LAG
45	814	84.9	180	3	AAW70860	Human LAG

ALIGNMENTS

RESULT 1	
AAW62584	
ID	AAW62584 standard; protein; 180 AA.
XX	
AC	AAW62584;
XX	
DT	17-SEP-1998 (first entry)
XX	
DE	Cancer associated antigen NY-ESO-1.
XX	
KW	Cancer associated antigen; NY-ESO-1; regression; progression; onset;
KW	cancer; treatment; diagnosis.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 7
FT	Misc-difference 9 /note= "potential myristoylation site"
FT	Misc-difference 11 /note= "potential myristoylation site"
FT	Misc-difference 98 /note= "potential phosphorylation site"
FT	Misc-difference 134 /note= "potential phosphorylation site"
FT	Misc-difference 138 /note= "potential phosphorylation site"
XX	WO9814464-A1.
XX	
PD	09-APR-1998.
XX	
PF	15-SEP-1997; 97WO-US016335.
XX	
PR	03-OCT-1996; 96US-00725182.
XX	
PA	(LUDW-) LUDWIG INST CANCER RES.
XX	
PI	Chen Y, Scanlan M, Gure A, Old LJ, Jager E, Knuth A;
PI	Driffoht JW;
XX	
DR	WPI; 1998-286417/25.
XX	
DR	N-PSDB; AAV38566.
XX	
PT	New isolated cancer associated antigen - is used to develop products for the diagnosis and treatment of cancers and for monitoring cancer therapy.

```
XX Claim 8; Fig 3; 49pp; English.
XX
CC The present sequence represents a cancer associated antigen. The clone
CC from which the DNA sequence is obtained is designated NY-ESO-1. The
CC specification described a method for determining regression, progression
CC of onset of a cancerous condition, comprising monitoring a sample from a
CC patient with the cancerous condition for a parameter selected from NY-ESO
CC -1 protein, a peptide derived from NY-ESO-1 protein and cytolytic T cells
CC specific for the peptide and an MHC molecule with which it non-covalently
CC complexes. Methods for the treatment of a cancerous condition are also
CC described. The NY-ESO-1 protein and peptides derived from it can be used
CC for diagnosis and treatment of cancers and to monitor the efficacy of a
CC therapeutic regime
XX
XX SQ Sequence 180 AA;
    Query Match      100.0%; Score 959; DB 2; Length 180;
    Best Local Similarity 100.0%; Pred. No. 5.9e-79;
    Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGSGTGDADGPGGPGIPDPGPGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
DB 1 MQAEGRTGSGTGDADGPGGPGIPDPGPGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLGCGCCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLGCGCCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
RESULT 2
AAW69665
ID AAW69665 standard; protein; 180 AA.
AC AAW69665;
XX
XX 27-OCT-1998 (first entry)
XX Human NY-ESO-1 protein sequence, formerly known as LL-1.2 clone.
XX Human; LAGE-1; tumour associated protein; LL-1; diagnosis; tumour.
XX
XX Homo sapiens.
XX WO9832855-A1.
XX 30-JUL-1998.
XX
XX 27-JAN-1998; 98WO-US001445.
XX
XX 27-JAN-1997; 97US-00791495.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Lethe B, Lucas S, De Smet C, Godelaine D, Boon-Falleur T;
XX WPI; 1998-427951/36.
XX N-PSDB; AAV50348.
XX
XX New isolated LAGE-1 tumour associated nucleic acids - used to develop
XX products for the diagnosis and treatment of LAGE-1 associated disorders,
XX particularly tumours.
XX
XX Example 2; Page 57-58; 73pp; English.
XX
XX The present sequence represents human NY-ESO-1, formerly known as LL-1.2
XX clone, which is used in an example from the present invention which
XX describes LAGE-1 tumour associated protein (TAP). The present invention
XX also describes: (1) a method for treating a subject with a disorder
CC
```

```
CC characterised by expression of a LAGE-1 nucleic acid molecule or an
CC expression product, comprising administering to the subject autologous
CC cytolytic T cells to ameliorate the disorder, where the cytolytic T cells
CC are specific for complexes of an HLA molecule and a LAGE-1 TAP or an
CC immunogenic fragment; (2) a method for treating a subject with a disorder
CC characterised by expression of a LAGE-1 nucleic acid molecule or an
CC expression product, comprising administering a LAGE-1 TAP or an
CC immunogenic fragment to ameliorate the disorder; and (3) a method for
CC selectively enriching a population of T cells with cytolytic T cells
CC specific for a LAGE-1 TAP comprising contacting an isolated population of
CC T cells with an agent presenting a complex of a LAGE TAP or an
CC immunogenic fragment and a HLA presenting molecule to selectively enrich
CC the isolated population of T cells with the cytolytic T cells. The
CC methods and products from the present invention can be used for the
CC diagnosis and treatment of LAGE-1 associated disorders, particularly
CC tumours
XX
XX SQ Sequence 180 AA;
    Query Match      100.0%; Score 959; DB 2; Length 180;
    Best Local Similarity 100.0%; Pred. No. 5.9e-79;
    Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGSGTGDADGPGGPGIPDPGPGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
DB 1 MQAEGRTGSGTGDADGPGGPGIPDPGPGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLGCGCCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLGCGCCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
RESULT 3
AAW05965
ID AAW05965 standard; protein; 180 AA.
XX
XX AAW05965;
XX
XX 16-AUG-1999 (first entry)
XX Human cancer antigen NY ESO-1/CAG-3 ORF1 protein.
XX
XX NY ESO-1/CAG-3 gene; CAG-3 gene; cancer antigen; human; leukaemia;
XX non-Hodgkins lymphoma; Hodgkins lymphoma; lung cancer; metastasis;
XX melanoma; adenocarcinoma; thymoma; colon cancer; uterine cancer;
XX breast cancer; prostate cancer; ovarian cancer; cervical cancer;
XX bladder cancer; kidney cancer; pancreatic cancer; liver cancer; sarcoma;
XX tumour; diagnosis; immunotherapy; therapy; vaccine; ORF1.
XX
XX Homo sapiens.
XX
XX WO9918206-A2.
XX
XX 15-APR-1999.
XX
XX 21-SEP-1998; 98WO-US019609.
XX
XX 08-OCT-1997; 97US-0061428P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Wang RF, Rosenberg SA;
XX WPI; 1999-277270/23.
XX N-PSDB; AAX58599.
XX
XX Cancer antigen NY ESO1/CAG-3.
XX Claim 4; Fig 3A; 88pp; English.
XX
```

XX The present sequence represents the ORF1 protein encoded by open reading
CC frame 1 of the human ESO-1/CAG-3 (or CAG-3) gene. CAG-3 is a new and
CC potent tumour antigen capable of eliciting an antigen specific immune
CC response by T cells. Cancer peptides comprising ORF1, ORF2 (see
CC AAY05966), portions of these peptides and their variants (see AAY05965-
CC 87), are useful as cancer vaccines that protect the recipient from
CC development of cancer. The invention provides: vectors and host cells
CC (also useful as vaccines); a method of diagnosis of cancer or precancer;
CC a transgenic animal; antisense oligonucleotides that inhibit expression
CC of the cancer peptide or tumour antigen; antibodies reacting with the CAG
CC -3 cancer peptide, useful in diagnostic and detection assays; and methods
CC for preventing or inhibiting cancer by administering a cancer peptide,
CC with or without an HLA molecule. The cancer peptides form part of, or are
CC derived from, cancers such as primary or metastatic melanoma, thymoma,
CC lymphoma, sarcoma, lung cancer, liver cancer, leukaemia, uterine cancer,
CC cervical cancer, bladder cancer, kidney cancer and adenocarcinomas such
CC as breast, prostate, ovarian, pancreatic and thyroid cancers. Melanoma is
CC treated by inducing cancer-specific T cells in vitro for subsequent
CC return to a patient
XX
SQ Sequence 180 AA;

Query Match 100.0%; Score 959; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGSGTDADGPGGPGTDPGCGNAGPGGAGATGCGPRGAGNARASGPGGA 60
Db |||||
QY 1 MQAEGRTGSGTDADGPGGPGTDPGCGNAGPGGAGATGCGPRGAGNARASGPGGA 60
Db |||||
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLSFYLLAMPATPWEALRRSLAQDAPPLVPFG 120
Db |||||
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLSFYLLAMPATPWEALRRSLAQDAPPLVPFG 120
Db |||||
QY 121 VLLKEFTVSGNLTIRLTAAHRLQLSISCLQLSLLMWITQCPLPVLQPPSGQR 180
Db |||||
QY 121 VLLKEFTVSGNLTIRLTAAHRLQLSISCLQLSLLMWITQCPLPVLQPPSGQR 180
Db |||||

RESULT 4
AAY52430
ID AAY52430 standard; protein; 180 AA.

XX AAY52430;
XX
XX
XX
DT 21-OCT-2004 (revised)
DT 15-FEB-2000 (first entry)
XX
XX Human tumour antigen NY-ESO-1.
XX
XX Cancer; tumour; antigen; MHC; major histocompatibility complex; T-cell;
KW cytotoxic; helper; stimulation; proliferation; treatment; diagnosis;
KW prevention; melanoma; breast cancer; ovarian cancer; prostate cancer;
KW hepatoma; thyroid cancer; bladder cancer; lung cancer; lymphoma.
XX
XX Homo sapiens.
OS Unidentified.
XX

Key Location/Qualifiers
FH Peptide 44..53
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 60..69
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 60..68
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 63..72
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 79..88
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 79..87
FT /note= "Peptide presented by MHC Class I HLA-B7, HLA-B8
FT and HLA-B35"

FT Peptide 82..91
FT /note= "Peptide presented by MHC Class I HLA-A1"
FT Peptide 82..90
FT /note= "Peptide presented by MHC Class I HLA-A1"
FT Peptide 83..91
FT /note= "Peptide presented by MHC Class I HLA-B44"
FT Peptide 84..92
FT /note= "Peptide presented by MHC Class I HLA-B7, HLA-B8
FT and HLA-B35"
FT Peptide 87..96
FT /note= "Peptide presented by MHC Class I HLA-A1"
FT Peptide 88..96
FT /note= "Peptide presented by MHC Class I HLA-B44"
FT Peptide 96..104
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 100..108
FT /note= "Peptide presented by MHC Class I HLA-B44"
FT Peptide 102..110
FT /note= "Peptide presented by MHC Class I HLA-B44"
FT Peptide 107..116
FT /note= "Peptide presented by MHC Class I HLA-A24"
FT Peptide 110..118
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT Peptide 113..122
FT /note= "Peptide presented by MHC Class I HLA-B7 and HLA-
FT B52"
FT Peptide 113..121
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 115..124
FT /note= "Peptide presented by MHC Class I HLA-A3"
FT Peptide 118..126
FT /note= "Peptide presented by MHC Class I HLA-B35"
FT Peptide 124..133
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT Peptide 125..133
FT /note= "Peptide presented by MHC Class I HLA-A24"
FT Peptide 138..147
FT /note= "Peptide presented by MHC Class I HLA-B8"
FT Peptide 139..147
FT /note= "Peptide presented by MHC Class I HLA-B7"
FT Peptide 145..153
FT /note= "Peptide presented by MHC Class I HLA-A24 and HLA-
FT B52"
FT Peptide 153..162
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT Peptide 154..163
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT Peptide 154..162
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT Peptide 156..167
FT /note= "Peptide (AAY52434) presented by MHC Class I HLA-
FT A2"
FT Peptide 158..166
FT /note= "Peptide presented by MHC Class I HLA-A3"
FT Peptide 159..167
FT /note= "Peptide presented by MHC Class I HLA-A3"
FT Peptide 162..170
FT /note= "Peptide presented by MHC Class I HLA-B52"
FT
XX WO953938-A1.
XX
XX 28-OCT-1999.
XX
XX 24-MAR-1999; 99WO-US006875.
XX
XX 17-APR-1998; 98US-00062422.
XX 02-OCT-1998; 98US-00165546.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Stockert E, Jager E, Chen Y, Scanlan M, Alexander K, Old LJ;
PI Gure A, Ritter G;
XX

```
DR WPI; 2000-038483/03.
DR N-PSDB; AAZ38380.
XX
XX Novel peptides which bind to MHC class I and MHC class II molecules,
XX useful for therapeutic and diagnostic purposes.
XX
XX Claim 30; Fig 3; 49pp; English.
XX
XX This sequence represents a human tumour antigen, NY-ESO-1, the cDNA
XX encoding which was isolated from an oesophagus squamous cell cancer cDNA
XX library. Tissue localisation studies revealed it to be expressed at high
XX levels in normal ovary and testis but not in normal colon, kidney, liver,
XX brain, oesophagus and skin. It was expressed in certain tumours and
XX tumour cell lines with some degree of frequency - these included melanoma
XX specimens and cell lines, and breast and bladder cancer specimens, with
XX expression in other tumour types being sporadic. Peptides derived from NY
XX -ESO-1 are bound by both MHC (major histocompatibility complex) Class I
XX and Class II molecules for presentation to T-cells. Peptides AAY52431-
XX Y52434 bind to Class I HLA-A2 molecules, thereby stimulating
XX proliferation of cytotoxic T-cells, while peptides AAY52435-Y52440 bind
XX to Class II HLA-DR3 molecules, stimulating helper T-cell proliferation.
XX The peptides derived from NY-ESO-1 may be used in methods and
XX compositions used for the treatment, diagnosis and prevention of cancers
XX (such as melanoma, breast cancer, prostate cancer, lung cancer, hepatoma,
XX ovarian cancer, thyroid cancer, bladder cancer, or lymphoma) and to
XX stimulate the proliferation of T cells
XX
XX Revised record issued on 21-OCT-2004 : Correction to feature table key
XX
XX Sequence 180 AA;
XX
XX Query Match 100.0%; Score 959; DB 3; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-79;
XX Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQAEGRTGSGTGDADGGPGGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADGGPGGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
RESULT 5
AAY70862
ID AAY70862 standard; protein; 180 AA.
XX
XX AAY70862;
XX
XX 31-JUL-2000 (first entry)
XX
XX Human tumour antigen, NY-ESO-1 protein.
XX
XX NY-ESO-1; CAMEL; CTL-recognised Antigen on MELANOMA; human; cancer; CTL;
XX cytotoxic T lymphocyte; tumour-associated antigen; LAGE-1; anticancer;
XX melanoma; immunotherapy; immune response.
XX
XX Homo sapiens.
XX
XX W0200023584-A1.
XX
XX 27-APR-2000.
XX
XX 15-OCT-1999; 99WO-EP007832.
XX
XX 16-OCT-1998; 98EP-00119583.
XX
XX (BOEH ) BOEHRINGER INGELHEIM INT GMBH.
XX
PA
XX
XX (UYHO-) UNIV HOSPITAL LEIDEN.
XX
XX Schrier PI, Aarnoudse CA, Heider K, Klade C;
XX
XX WPI; 2000-339685/29.
XX
XX N-PSDB; AAD00152.
XX
XX Tumor-associated antigen useful for cancer immunotherapy is encoded by
XX the open reading frame of LAGE-1 (a tumor-specific antigen) cDNA.
XX
XX Example 3; Page 62-63; 73pp; English.
XX
XX The present sequence is the human NY-ESO-1 protein, a tumour antigen,
XX identified by screening an esophagus carcinoma cDNA library. This protein
XX is derived from open reading frame (ORF)-1 that contain epitopes of
XX tumour specific T-cells. NY-ESO-1 is expressed in different tumour types,
XX but not in healthy tissues except in testis. It also shows homology with
XX the CAMEL (Cytotoxic T lymphocytes (CTL)-recognised Antigen on MELANOMA)
XX protein, a tumour-associated antigen. The tumour-associated antigen
XX This sequence has anticancer activity. CAMEL tumour antigen and
XX immunogenic peptides derived from it are useful for cancer immunotherapy.
XX They have the potential to induce an immune response, by eliciting a CTL
XX response. The DNA molecule is used for the construction of recombinant or
XX fusion proteins
XX
XX Sequence 180 AA;
XX
XX Query Match 100.0%; Score 959; DB 3; Length 180;
XX Best Local Similarity 100.0%; Pred. No. 5.9e-79;
XX Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQAEGRTGSGTGDADGGPGGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
Db 1 MQAEGRTGSGTGDADGGPGGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMLWITQCFLPVFLAQPSPGQRR 180
RESULT 6
AAB03154
ID AAB03154 standard; protein; 180 AA.
XX
XX AAB03154;
XX
XX 23-OCT-2000 (first entry)
XX
XX Human oesophageal cancer-associated antigen NY-ESO-1.
XX
XX Oesophageal cancer associated antigen; NY-ESO-1; human; immunogen;
XX oesophageal carcinoma; melanoma; ovary; testis; transmembrane domain;
XX antibody; diagnostic marker; drug delivery target.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 7
XX /note= "Potential N-myristoylation site"
XX Modified-site 9
XX /note= "Potential N-myristoylation site"
XX Modified-site 11
XX /note= "Potential O-phosphorylation site"
XX Modified-site 98
XX /note= "Potential O-phosphorylation site"
XX Modified-site 134
XX /note= "Potential O-phosphorylation site"
XX Modified-site 138
XX /note= "Potential O-phosphorylation site"
```



```
AAE07714 standard; protein; 180 AA.
AAE07714;
06-NOV-2001 (first entry)
Human NY ESO-1 protein.
Human; major histocompatibility complex; MHC; vaccine; metastasis;
class II restricted T cell epitope; MHC-II epitope; cancer antigen;
NY ESO-1 protein; CD4+ T lymphocyte; human leucocyte antigen; HLA;
tumour-specific humoral-mediated immunity; cancer; cytostatic;
immunotherapy.
Homo sapiens.
Key Location/Qualifiers
Misc-difference 45..47 /note="Encoded by CCCGGGGC"
WO200155393-A2.
02-AUG-2001.
26-JAN-2001; 2001WO-US002765.
28-JAN-2000; 2000US-0179004P.
29-SEP-2000; 2000US-0237107P.
(USSH ) US DEPT HEALTH & HUMAN SERVICES.
Wang R, Rosenberg SA, Zeng G;
WPI: 2001-496851/54.
N-PSDB; AAD14179, AAD14180.
New NY-ESO cancer peptide or MHC class II restricted T cell epitopes,
useful as immunogen and vaccine for inhibiting cancer in a mammal or as
protection from metastasis.
Example 1; Fig 1; 134pp; English.
The invention relates to the identification and isolation of major
histocompatibility (MHC) class II restricted T cell epitope (MHC-II
epitope) derived from the cancer antigen, NY ESO-1. The MHC-II epitopes
from NY ESO-1 are recognised by CD4+ T lymphocytes in an human leucocyte
antigen (HLA) class II restricted manner, in particular HLA-DR or HLA-DP
restricted. The products of the gene are promising candidates for
immunotherapeutic strategies for the prevention, treatment and diagnosis
of patients with cancer. The cancer epitopes are useful as immunogen and
vaccine to inhibit or to prevent cancer in a mammal by eliciting CD4+ T
lymphocytes resulting in protection of the recipient from development of
cancer and protection from metastasis, or by inhibiting the growth of
cells expressing the NY-ESO-1 gene product. The cancer peptides are also
useful as diagnostic agent to detect the presence of cancer, to enhance
the generation of antibody and/or CD8+ T cell responses against any given
target antigen and/or hapten and to induce tumour-specific humoral-
mediated immunity against cancer. The present sequence is human NY ESO-1
protein
Sequence 180 AA;
Query Match 100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGCGRPGAGAAASGPGGA 60
DB 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGCGRPGAGAAASGPGGA 60
QY 61 PRGPHGASGLNGCCRCGARGPESLLEFFYLAMPATPWEALRRSLAQADPLVPVG 120
DB 61 PRGPHGASGLNGCCRCGARGPESLLEFFYLAMPATPWEALRRSLAQADPLVPVG 120
```

```
121 VLLKEFTVSGNIIITRLTAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQR 180
121 VLLKEFTVSGNIIITRLTAADHRQLQLSISSCLQQLSLLMWITQCFLPVFLAQPSPGQR 180

RESULT 11
AAU84818
ID AAU84818 standard; protein; 180 AA.
XX
AC AAU84818;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human NYNSO1a consensus sequence.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Homo sapiens.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU ) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
WPI: 2002-147575/19.
XX
New synthetic polypeptides having several different segments of at least
one parent polypeptide linked together differently compared to the
linkage in the parent polypeptide, for inducing immune response against a
pathogen or cancer.
Example 3; Fig 27; 364pp; English.
The invention relates to a new synthetic polypeptide (I) comprising
several different segments of at least one parent polypeptide linked
together in a different relationship relative to their linkage in the
parent polypeptide to impede, abrogate or otherwise alter at least one
function associated with the parent polypeptide and for inducing an
immune response against a pathogen or cancer. Also included are a
synthetic polynucleotide encoding and a computer system for designing the
synthetic polypeptides. The synthetic polypeptides and polynucleotides
are referred to as a Savine. The synthetic polypeptide is useful for
modulating immune responses preferably directed against a pathogen or a
cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
oesophagus, brain, testicle, uterus), as potentiating agents.
Compositions comprising the polypeptide may be used in the treatment or
prophylaxis against viral (such as infections caused by HIV (human
immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
(e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
a consensus sequence for a parent protein used to design a savine of the
invention
Sequence 180 AA;
Query Match 100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


QY 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
DB 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFFYLAMPFATPMEALARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFFYLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQLQLSLLMWITQCFLPVFLAQPPSGQRR 180
DB 121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQLQLSLLMWITQCFLPVFLAQPPSGQRR 180
RESULT 12
AAU11543
IL AAU11543 standard; protein; 180 AA.
XX
XX AAU11543;
XX AC
XX 12-MAR-2002 (first entry)
XX DE Human tumour associated antigen NY-ESO.
XX DE
XX Human; tumour associated antigen; NY-ESO; human leukocyte antigen;
KW major histocompatibility complex; MHC; HLA-A2.2; vaccine; cancer; HIV;
KW human immunodeficiency virus infection; cytostatic; virucide;
KW housekeeping epitope; adoptive immunotherapy; neoplastic disease;
KW viral disease; hepatitis virus; papilloma virus; tumour; leukaemia;
KW lymphoma; breast cancer; prostate cancer; lung cancer;
KW parasitic infection; Chlamydia; Trypanosoma; Toxoplasma.
XX
XX Homo sapiens.
XX
XX W0200182963-A2.
XX PN
XX 08-NOV-2001.
XX PD
XX 27-APR-2001; 2001WO-US013806.
XX PP
XX 28-APR-2000; 2000US-00560465.
XX PR 28-APR-2000; 2000US-00561074.
XX PR 28-APR-2000; 2000US-00561571.
XX PR 28-APR-2000; 2000US-00561572.
XX
XX (CTL1-) CTL IMMUNOTHERAPIES CORP.
XX PA
XX Simard JUL, Diamond DC, Lei X;
XX P:
XX WPI; 2002-0666492/09.
XX DR
XX
XX Novel vaccine useful for treating neoplastic and viral diseases,
P2 comprises a first housekeeping epitope derived from a first antigen
P1 associated with a first target cell.
XX
XX Example 23; Fig 17; 131pp; English.
XX P3
XX The invention relates to a vaccine comprising a first housekeeping
CC epitope derived from a first antigen associated with a first target cell.
CC Also included are an isolated T cell expressing a T cell receptor
CC specific for a major histocompatibility complex (MHC)-peptide complex
CC comprising a first housekeeping epitope which is derived from a first
CC antigen associated with a first target cell, selecting an epitope (or
CC peptide sequence) from a population of peptide fragments of an antigen
CC associated with a target in a host, where the fragments have a known or
CC predicted affinity for a MHC receptor peptide binding cleft of the host,
CC where the epitope selected corresponds to a product of proteolytic
CC cleavage of the antigen in a cell of the host and a nucleic acid
CC construct comprising a first coding region, where the first coding region
CC comprises a first sequence encoding at least a first polypeptide, where
CC the first polypeptide comprises a first housekeeping epitope derived from
CC a first antigen associated with a first target cell; The epitopes,
CC peptides, vaccines and nucleic acids are useful in the manufacture of a

CC medicament for use in adoptive immunotherapy and for prevention and
CC treatment of neoplastic and viral diseases (e.g. human immunodeficiency
CC virus, HIV infection, hepatitis virus and papilloma virus), cancers
CC (e.g. tumours, leukaemia, lymphoma, breast cancer, prostate cancer and
CC lung cancer), infection of cells by intracellular parasites (e.g.
CC Chlamydia, Trypanosoma and Toxoplasma) and many other examples given in
CC the specification. The invention permits the vaccine designer to ignore
CC peptides that, despite predicted high binding affinity for MHC, will
CC never be useful because they cannot be presented by target cells. The
CC invention provides a major advance in vaccine design, one that combines
CC the power of antigen sequence analysis with the fundamental realities of
CC immunology. The invention allows for the simple and effective selection
CC of meaningful epitopes for creation of MHC class I or Class II vaccines
CC using any polypeptide sequence corresponding to a desired target. The
CC present sequence is an HLA-A2.1 (human leukocyte antigen) presenting
CC target cell protein from which epitopes of the invention may be derived,
CC NY-ESO (a tumour associated antigen)
XX
SQ Sequence 180 AA;
Query Match 100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 5,9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
DB 1 MQAEGRTGSGTGDADPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFFYLAMPFATPMEALARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFFYLAMPFATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQLQLSLLMWITQCFLPVFLAQPPSGQRR 180
DB 121 VLLKEFTVSGNLTITRLTAADHRQLQLSISSCLQLQLSLLMWITQCFLPVFLAQPPSGQRR 180
RESULT 13
ABR58672
ID ABR58672 standard; protein; 180 AA.
XX
XX ABR58672;
XX AC
XX 09-JUL-2003 (first entry)
XX DT
XX Human cancer related protein SEQ ID NO:329.
XX DE
XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometriosis.
XX KW
XX Homo sapiens.
XX OS
XX WO2003025138-A2.
XX PN
XX 27-MAR-2003.
XX PD
XX 17-SEP-2002; 2002WO-US029560.
XX PF
XX 17-SEP-2001; 2001US-0323469P.
XX PR 20-SEP-2001; 2001US-0323887P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 08-FEB-2002; 2001US-0355145P.
XX PR 08-FEB-2002; 2002US-0355257P.
XX PR 12-APR-2002; 2002US-0372246P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX
XX WPI; 2003-354600/33.
XX DR N-PSDB; ACC72823.
XX

PT New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
PS Claim 12; Page 757-758; 767pp; English.

XX The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 180 AA;

Query Match 100.0%; Score 959; DB 6; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
DB 1 MOAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEALARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 14
ABR48210
ID ABR48210 standard; protein; 180 AA.
AC ABR48210;

DT 12-JUN-2003 (first entry)

DE Human bladder cancer associated protein sequence SEQ ID NO:139.

XX Human; bladder cancer; cytostatic; gene therapy; vaccine.

XX Homo sapiens.

OS WO2003003906-A2.

XX 16-JAN-2003.

XX 03-JUL-2002; 2002WO-US021338.

XX 03-JUL-2001; 2001US-0302814P.

PR 03-AUG-2001; 2001US-0310099P.

PR 08-NOV-2001; 2001US-0343705P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-037246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Mack DH, Aziz N;
XX WPI; 2003-201532/19.
DR N-PSDB; ACC51024.

XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.

PS Claim 10; Page 278; 307pp; English.

XX The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications

XX Sequence 180 AA;

Query Match 100.0%; Score 959; DB 6; Length 180;
Best Local Similarity 100.0%; Pred. No. 5.9e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
DB 1 MOAEGRTGCGTGADGPGGPGIPDGGNAGGPGGAGATGCGPRGAGAAASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEALARRSLAQDAPPLVPVG 120
DB 61 PRGPHGGAASGLNGCCCGARGPESRLLEFYLLAMPATPMEALARRSLAQDAPPLVPVG 120
QY 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 15
ABU56508
ID ABU56508 standard; protein; 180 AA.

AC ABU56508;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #101.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;

KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;

KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;

KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;

KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

OS WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

[illegible]

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:13:27 ; Search time 42 Seconds
(without alignments)
412.358 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRTGGTGADGPGG.....WITQCFLPVFLAQPFGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174.5	18.2	627	T35608	polyketide hydroxy
2	155.5	16.2	145	T08435	la costa protein
3	155.5	16.2	641	Q0B831	nuclear antigen EB
4	153	16.0	1079	D70807	hypothetical glyci
5	152.5	15.9	779	1 CGB01S	collagen alpha 1(I
6	151.5	15.8	1464	1 CGH01S	collagen alpha 1(I
7	151	15.7	1489	D70807	hypothetical glyci
8	151	15.7	1901	D70806	hypothetical glyci
9	150.5	15.7	530	D70812	hypothetical prote
10	149.5	15.6	615	D70589	hypothetical glyci
11	149.5	15.6	1215	D72734	myosin-IA - Acanth
12	149	15.5	297	D72525	hypothetical prote
13	149	15.5	299	D70605	hypothetical prote
14	148.5	15.5	252	1 S01821	glycine-rich prote
15	148.5	15.5	1042	1 CGH01S	collagen alpha 1(I
16	147.5	15.4	473	D70629	collagen - chicken
17	147	15.3	296	D721070	hypothetical prote
18	147	15.3	440	D70806	probable tail comp
19	147	15.3	1391	D70806	hypothetical glyci
20	146.5	15.3	290	1 CGH01S	hypothetical prote
21	146.5	15.3	1806	1 CGH01S	collagen alpha 1(I
22	146	15.2	338	1 K0MU	glycine-rich cell
23	146	15.2	853	D70896	hypothetical glyci
24	146	15.2	1880	D718531	tractin - mediana
25	145.5	15.2	228	D718531	glycine-rich prote
26	145.5	15.2	482	D718531	collagen alpha 1(I
27	145.5	15.2	1306	D70934	hypothetical glyci
28	145	15.1	741	D70917	hypothetical glyci
29	145	15.1	1329	D70917	hypothetical glyci

30	144.5	15.1	290	2 T24586	hypothetical prote
31	144.5	15.1	301	2 B31219	collagen 2 - Caeno
32	144.5	15.1	330	2 T26004	hypothetical prote
33	144.5	15.1	438	2 D90734	probable tail fibe
34	144.5	15.1	1147	1 MWAXIB	myosin heavy chain
35	144	15.0	667	2 A70893	hypothetical glyci
36	143.5	15.0	310	2 I50696	collagen alpha 1(I
37	143.5	15.0	314	2 T32247	hypothetical prote
38	143.5	15.0	714	2 A70807	hypothetical glyci
39	143.5	15.0	886	2 I50694	collagen alpha 1(I
40	143.5	15.0	1049	1 CGH07S	collagen alpha 1(I
41	143	14.9	323	2 T19142	hypothetical prote
42	143	14.9	437	2 H90854	probable tail fibe
43	143	14.9	439	2 C90769	probable tail fibe
44	143	14.9	1660	2 A70869	hypothetical glyci
45	142.5	14.9	244	2 T49893	glycine-rich prote

ALIGNMENTS

RESULT 1

T35608

polyketide hydroxylase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C:Accession: T35608

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL data Library, June 1999

A:Reference number: Z21584

A:Accession: T35608

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-627 <SEE>

A:Cross-references: UNIPROT:P42534; UNIPARC:UPI000012CFBE; EMBL:AL079356; PIDN:CAB45603.1

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC6G9.12c

C:Superfamily: tetracycline 6-hydroxylase

Query Match 18.2%; Score 174.5; DB 2; Length 627;
Best Local Similarity 42.7%; Pred. No. 4.1e-06;
Matches 47; Conservative 4; Mismatches 28; Indels 31; Gaps 6;

QY	5	GRGTGSGTGDADGPGGPGIPDGGNAGGPGAGATGG-RGPRGAGAAASGPGGAPRG	63
DB	409	GRGTGG-PGGPGGGLGGPGGPGG-TGGPGGPGGPGGPGGAGAPGGPGG----	462

QY	64	PHGGAASGLNGCCRCGARGPSRLLEFYLAMPF-----ATPMEAE	103
DB	463	PGGG-----GPGGILNALGYRYPGAVVGADPATVPVPE	498

RESULT 2

T08435

la costa protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004

C:Accession: T08435

R:Maleszka, R.; de Couet, H.G.; Miklos, G.L.

Proc. Natl. Acad. Sci. U.S.A. 95, 3731-3736, 1998

A:Title: Data transferability from model organisms to human beings: insights from the fur

A:Reference number: Z16415; MUID:98188272; PMID:9520435

A:Accession: T08435

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-145 <MAL>

A:Cross-references: UNIPROT:O61351; UNIPARC:UPI0000076890; EMBL:AF017777; NID:g3004652; I

A:Experimental source: strain Cantons

C:Genetics:

A:Gene: lcs

A:Cross-references: FlyBase:FBgn0024251

A;Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen chain operative melting of intact type I collagen.
A;Reference number: I55342; MUID:92042092; PMID:1718984
A;Accession: I55342
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 258-268;1347-1357 <DEA>
A;Cross-references: UNIPARC:UPI0000173B42; UNIPARC:UPI0000173B43; GB:S67495; NID:g239007
A;Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg reported
R;Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.
J. Biol. Chem. 245, 5042-5048, 1970
A;Title: Comparative study of glycopeptides derived from selected vertebrate collagens.
A;Reference number: A92069; MUID:71001508; PMID:4319110
A;Accession: A92069
A;Molecule type: protein
A;Residues: 263-268 <MOR>
A;Cross-references: UNIPARC:UPI000014DF11
A;Experimental source: skin
A;Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine
R;Labhard, M.E.; Hollister, D.W.
Matrix 10, 124-130, 1990
A;Title: Segmental amplification of the entire helical and telopeptide regions of the collagen type I alpha 1(I) chain
A;Reference number: S15989; MUID:90326017; PMID:2374517
A;Accession: S15989
A;Molecule type: mRNA
A;Residues: 281-302;402-420;823-843;925-944;1026-1045;1143-1162 <LAB>
A;Cross-references: UNIPARC:UPI0000173B44; UNIPARC:UPI0000173B45; UNIPARC:UPI0000173B46;
R;Wirtz, M.K.; Rao, V.H.; Glanville, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries, W.N.
Connect. Tissue Res. 29, 1-11, 1993
A;Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain of collagen type I
A;Reference number: I52905; MUID:93339042; PMID:8339541
A;Accession: I52905
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 342-352; 'C', 354-359 <WID>
A;Cross-references: UNIPARC:UPI000070EB6; GB:S64717; NID:g408195; PIDN:AAB27677.1; PID:
A;Note: mutant sequence from patient with osteogenesis imperfecta
R;Bernard, M.P.; Chu, M.L.; Myers, J.C.; Ramirez, F.; Eikenberry, E.F.; Prockop, D.J.
Biochemistry 22, 5213-5223, 1983
A;Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalpha 1(I) chain of collagen type I
A;Reference number: A90476; MUID:84080385; PMID:6689127
A;Accession: A90476
A;Molecule type: mRNA
A;Residues: 425-1250, 'S', 1330-1390, 'X', 1392-1464 <BER>
A;Cross-references: UNIPARC:UPI0000173B4A; GB:K01228; NID:g180391; PIDN:AAA51995.1; PID:
A;Note: sequence partially completed for missing nucleotides by A29439
R;Chu, M.L.; Gargiulo, V.; Williams, C.J.; Ramirez, F.
J. Biol. Chem. 260, 691-694, 1985
A;Title: Multixon deletion in an osteogenesis imperfecta variant with increased type I collagen
A;Reference number: A22161; MUID:85104934; PMID:2981843
A;Accession: A22161
A;Molecule type: DNA
A;Residues: 472-594, 'R', 596-607 <CH3>
A;Cross-references: UNIPARC:UPI000011F786; GB:K03178; GB:K03179; NID:g179612; NID:g17961
A;Note: the authors translated the codon CGT for residue 595 as Pro
R;Wallis, G.A.; Starman, B.J.; Zinn, A.B.; Byers, P.H.
Am J. Hum. Genet. 46, 1034-1040, 1990
A;Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained by a mutation in the alpha 1(I) chain of collagen type I
A;Reference number: A35336; MUID:90252792; PMID:23397700
A;Accession: A35336
A;Molecule type: mRNA
A;Residues: 710-720, 'E', 722-737, 'E', 739-745 <WAL>
A;Cross-references: UNIPARC:UPI0000173B4B
A;Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu
R;Forlino, A.; Zolezzi, F.; Valli, M.; Pignatti, P.F.; Cetta, G.; Brunelli, P.C.; Motte
Hum. Mol. Genet. 3, 2201-2206, 1994
A;Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the alpha 1(I) chain of collagen type I
A;Reference number: I54365; MUID:95187161; PMID:7881420
A;Accession: I54365
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 746-766, 'S', 768-781 <FOR>

A;Cross-references: UNIPARC:UPI000016A6FA; GB:L47667; NID:g1009093; PIDN:AAB59576.1; PID:
R;Chesler, S.D.; Wallis, G.A.; Byers, P.H.
J. Biol. Chem. 268, 18218-18225, 1993
A;Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of type I collagen
A;Reference number: A47426; MUID:93352846; PMID:8349697
A;Accession: A47426
A;Molecule type: mRNA
A;Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CHE>
A;Cross-references: UNIPARC:UPI000073A2A; GB:S64596; NID:g407589; PIDN:AAB27856.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBI:136445)
A;Note: does not represent an experimentally determined sequence but three different mutations
A;Accession: B47426
A;Molecule type: mRNA
A;Residues: 1179-1464 <CH4>
A;Cross-references: UNIPARC:UPI0000173B4C
A;Experimental source: normal dermal fibroblast culture
A;Accession: C47426
A;Molecule type: mRNA
A;Residues: 1179-1276, 'H', 1278-1464 <CH5>
A;Cross-references: UNIPARC:UPI0000173B4D
A;Experimental source: fetal cell 86-237
A;Accession: D47426
A;Molecule type: mRNA
A;Residues: 1179-1336, 1339-1464 <CH6>
A;Cross-references: UNIPARC:UPI0000173B4E
A;Experimental source: fetal cell 86-146
A;Accession: E47426
A;Molecule type: mRNA
A;Residues: 1179-1387, 'R', 1389-1464 <CH7>
A;Cross-references: UNIPARC:UPI0000173B4F
A;Experimental source: fetal cell 88-251
R;Cohn, D.H.; Apone, S.; Byre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.; Nict
J. Biol. Chem. 263, 14605-14607, 1988
Query Match 15.8%; Score 151.5; DB 1; Length 1464;
Best Local Similarity 37.2%; Pred. No. 0.00051;
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;
QY 3 ASGR-GTGGSTGDA--DGGPGGPIPDGPGNAGGPGGAGATGGRGPRGA---GAARASG 55
Db 825 AKGPGDAGKAGDAGPPGPGAGPGPIGNVGAQKAGSAGSGPGATGPGGAAGRVG 884
QY 56 P---GGGAPRPGHGAASGLNGCCRCGARGSPESRLLEFLYLPMPFATPMEALARRSLAQ 111
Db 885 PRGPGNAGPPGPPGPAG-----KEGKGPRGE-----TGPAGRPGEV----- 922
QY 112 DAPPLPVP 120
Db 923 -GPPGP-PG 929
RESULT 7
D70807
hypothetical glycine-rich protein RV3514 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: D70807
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sgares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70807
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1489 <COL>
A;Cross-references: UNIPARC:UPI00000D3AEB; GB:AL022022; GB:AL123456; NID:g3261554; PIDN:
C;Genetics:
A;Gene: RV3514
C;Superfamily: collagen alpha 1(IV) chain

```

Query Match          15.7%; Score 151; DB 2; Length 1489;
Best Local Similarity 41.1%; Pred. No. 0.00057;
Matches 37; Conservative 2; Mismatches 31; Indels 20; Gaps 3;

QY  5  GRGTGSGTGDADGPGGP-----GIPDGPGNAGGPGEA-----GATGGRGPRCA 48
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db   140 GAGAACAPGQAGGPGGAGLWNGNPGGAGSGGGGTGGAGGAGGWLFGVGGAGGVGAGG 199
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY  49  GAARASGPG----GGAPRGPHGGAASGLNG 74
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db   200 GTGGAGGPGGLIWGGGGAGGVGAGGGTGG 229
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
F70806
hypothetical glycine-rich protein Rv3508 - Mycobacterium tuberculosis (strain H37RV)
C.Species: Mycobacterium tuberculosis
C.Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C.Accession: F70806
R.;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A.Authors: Squares, R.; Sultston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A.Reference number: A70500; MUID:98295987; PMID:9634230
A.Accession: F70806
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-1901 <COL>
A.Cross-references: UNIPROT:O53553; UNIPARC:UPI000013C2A6; GB:AL022022; GB:AL123456; NID
A.Experimental source: strain H37RV
C.Genetics:
A.Gene: Rv3508
C.Superfamily: collagen alpha 1(IV) chain

```

```

Query Match      15.7%; Score 151; DB 2; Length 1901;
Best Local Similarity 41.1%; Pred. No. 0.00071;
Matches 37; Conservative 2; Mismatches 31; Indels 20; Gaps 3;

QY  5  GRGTGGSTGDADCGPGP-----GIPDGPNGAGGPGEA-----GATGGRGPRCA 48
DB  140 GAGAAGAPGQAGGPGGAGGPGAGLWNGGPGAGGSGGTGGAGGAGGWLFGVGGAGGVGGAGG 199

QY  49  GAARASGPG----GGAPRGPHPHGGAAASGLNG 74
DB  200 GTGGAGGPGGLIWGGGAGGAGGVGGAGGCTGG 229

RESULT 9
T32812
hypothetical protein H17B01.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32812
R:Gattung, S.; Maggi, L.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid H17B01.
A:Reference number: Z21227
A:Accession: T32812
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-530 <GAT>
A:Cross-references: UNIPROT:O61209; UNIPARC:UPI00000792BA; EMBL:AF040646; PIDN:AAB94986.
A:Experimental source: strain Bristol N2; clone H17B01
C:Genetics:
A:Gene: CESP:H17B01.2
A:Map position: 2
A:Introns: 42/3; 58/1; 173/3; 268/2; 308/2; 340/1; 364/2; 387/3

Query Match      15.7%; Score 150.5; DB 2; Length 530;
Best Local Similarity 32.3%; Pred. No. 0.00024;

```


Query Match 15.6%; Score 149.5; DB 2; Length 1215;
Best Local Similarity 41.0%; Pred. No. 0.00061;
Matches 41; Conservative 3; Mismatches 33; Indels 23; Gaps 4;

QY 5 GRGTGGS-----TGDADGPG---GPGIPDPGPGNAGGPGGAGATGGGPG 45
DB 995 GRGAGRALPMGGDEAPLPVPTGPGSGPGGPGAGPGRGMFAGPGRGPGGPGGPGPGP 1054

QY 46 RGAGAARASGPGGGAP---RGPHGGAASGLNGCCRCGARG 82
DB 1055 AGPGRG-APGPGRGAPGPGRGPGGPGGPPGGRGMPPPGRG 1093

RESULT 12
T27525
hypothetical protein ZC373.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T27525
R:Kershaw J.
submitted to the EMBL Data Library, April 1995
A:Reference number: Z20382
A:Accession: T27525
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-297 <WIL>
A:Cross-references: UNIPROT:Q23263; UNIPARC:UPI0000077A2E; EMBL:Z49131; PIDN:CAA88979.1;
A:Experimental source: clone ZC373
C:Genetics:
A:Gene: CRSP:ZC373.7
A:Map position: X
A:Introns: 118/1

Query Match 15.5%; Score 149; DB 2; Length 297;
Best Local Similarity 45.8%; Pred. No. 0.00019;
Matches 38; Conservative 2; Mismatches 33; Indels 10; Gaps 4;

QY 7 GTGGSTGTDADGPGGPGIPDPGPGN-----AGGPGGAGATGGRGPGCAAGARASG-PGGGA 60
DB 170 GRPQPQGSAGHPAQPGQSGKNNGAPGAAGGPGQPGRPKQGRKGSGAGAPGKAG 229

QY 61 PRGPHGGAASGLNGCCRCGARGP 83
DB 230 PAGPAGG--PGNNG--SAGTPGP 248

RESULT 13
T20605
hypothetical protein F08G5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20605
R:Steward, C.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z19299
A:Accession: T20605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-299 <WIL>
A:Cross-references: UNIPROT:Q19218; UNIPARC:UPI000007C4A1; EMBL:Z70682; PIDN:CAA94581.1;
A:Experimental source: clone F08G5
C:Genetics:
A:Gene: CRSP:F08G5.4
A:Map position: 4
A:Introns: 47/3

Query Match 15.5%; Score 149; DB 2; Length 299;
Best Local Similarity 40.4%; Pred. No. 0.00019;
Matches 36; Conservative 6; Mismatches 37; Indels 10; Gaps 3;

QY 7 GTGGSTGTDADGPGGPGIPDPGPGNAGGPGGAGATGGRGPGCAAGARASG-----PGGAPR 62
DB 160 GQPGQDGPAGQPGAPG--PAGPEGDAGAPGAGAPGAPGQDQGRGTGLPGAPGAPGPQ 218

QY 63 GPHG-----GAASGLNGCCRCGARGPESR 86
DB 219 GFSGNPGQDGAAGAPQAGPAGPAGPDQ 247

RESULT 14
S01821
glycine-rich protein 1.0 precursor - kidney bean
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 30-Sep-1989 #sequence_revision 19-May-1994 #text_change 09-Jul-2004
C:Accession: S01821
R:Keller, B.; Sauer, N.; Lamb, C.J.
EMBO J. 7, 3625-3633, 1988
A:Title: Glycine-rich cell wall proteins in bean: gene structure and association of the
A:Reference number: S01820; MUID:89091109; PMID:3208742
A:Accession: S01821
A:Molecule type: DNA
A:Residues: 1-252 <KEL>
A:Cross-references: UNIPROT:P10495; UNIPARC:UPI000012BAFE; EMBL:X13595; NID:g21000; PIDN:
C:Superfamily: Phaseolus glycine-rich protein 1.0
C:Keywords: cell wall; structural protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-252/Product: glycine-rich protein 1.0 #status predicted <MAT>

Query Match 15.5%; Score 148.5; DB 1; Length 252;
Best Local Similarity 44.4%; Pred. No. 0.00017;
Matches 36; Conservative 3; Mismatches 37; Indels 5; Gaps 2;

QY 4 EGRGTGGSTGTDADGPGGPGIPDPGPGNAGGPGGAGATGG--RGPRGAGARASGPGGGAP 61
DB 154 EGSGAGGYYGGANGGGGGGGGGGSGAAGGGAAGGAGGAGGAGGGYGGGAAGGGG 213

QY 62 RGPHGGAASGLNGCCRCGARG 82
DB 214 RSGSGGGGGYGG---GGARG 231

RESULT 15
CGCH15
collagen alpha 1(I) chain - chicken (tentative sequence) (fragments)
C:Species: Gallus gallus (chicken)
C:Date: 12-Aug-1981 #sequence_revision 06-Jul-1982 #text_change 31-Mar-2000
C:Accession: A90458; A90181; A02857
R:Highberger, J.H.; Corbett, C.; Dixit, S.N.; Yu, W.; Seyer, J.M.; Kang, A.H.; Gross, J.
Biochemistry 21, 2048-2055, 1982
A:Title: Amino acid sequence of chick skin collagen alpha1(I)-C88 and the complete primary
A:Reference number: A90458; MUID:82231995; PMID:7093229
A:Accession: A90458
A:Molecule type: protein
A:Residues: 1-1036 <HIG>
A:Cross-references: UNIPARC:UPI0000173B62
A:Experimental source: skin
A:Note: this is the latest in a series of papers from these workers elucidating the sequ
R:Eyre, D.R.; Glimcher, M.J.
Biochem. Biophys. Res. Commun. 48, 720-726, 1972
A:Title: Evidence for a previously undetected sequence at the carboxyterminus of the alp
A:Reference number: A90181; MUID:72243016; PMID:5047697
A:Accession: A90181
A:Molecule type: protein
A:Residues: 1037-1042 <EYR>
A:Cross-references: UNIPARC:UPI0000173B63
A:Experimental source: skin
A:Note: residues 1037-1042 above correspond to the carboxyl end of the protein
C:Comment: Lysines at positions 103, 700, 934, and 946 above may be hydroxylated in some
C:Comment: Most of the prolines at the third position of the tripeptide repeating unit
C:Comment: Pro-1002 is the only 3-hydroxyproline and the only hydroxylated proline in po
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer;
F:1/Modified site: pyrrolidine carboxylic acid (Gln) #status experimental

Query Match 15.5%; Score 148.5; DB 1; Length 1042;
Best Local Similarity 29.3%; Pred. No. 0.00063;

Matches		54;	Conservative	5;	Mismatches	50;	Indels	75;	Gaps	7;
Qy	6	RGTGGSTG	DADGPGPGIP	-----	DGPGNAGG	PGGAGATG	GRGPRG	-	47	
Db	871	KGRGETG	PAGPGAGPGAPG	PVGPAGKNG	DRGETG	PAGPGAGPG	PAGPGPGP	QGP	930	
Qy	48	-----	-----	-----	AGAAASG	PGGAPR	PHGGAAS	-	70	
Db	931	RGDKETG	EQDRGMKGHRG	FSGLQGP	PPPGAPGE	QPSGASG	PAG--	PRGPPGS	AGAA	988
Qy	71	---GLNCC-	CRCGARGP	ESRLLLE	FYLAMP	FPATPME	AEARRSLA	QDAPPL	PVPGVLLKEF	126
Db	989	GKDGLNGL	PGPIGPPG	PRGTGE	VGVGPPGP	-----	GPPGP-	EGPPSGG	F	1034
Qy	127	TVSG	130							
Db	1035	DLDG	1038							

Search completed: February 4, 2006, 12:18:10
Job time : 42 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:10:02 ; Search time 248 Seconds
(without alignments)
512.077 Million cell updates/sec

Title: US-09-529-206E-4
Perfect score: 959
Sequence: 1 MQAEGRGTCGTGADPGG.....WITQCFLPVFLAQPFGQRR 180

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	959	100.0	180	1 CTG1B_HUMAN	P78358 homo sapien
2	959	100.0	180	2 Q7LBY4_HUMAN	Q7LBY4 homo sapien
3	595	62.0	210	1 CTAG2_HUMAN	O75638 homo sapien
4	581	60.6	142	2 Q9NY13_HUMAN	Q9NY13 homo sapien
5	221	23.1	143	1 ITBA2_HUMAN	Q14657 homo sapien
6	222	23.1	143	2 Q5HY39_HUMAN	Q5HY39 homo sapien
7	174.5	18.2	627	1 HYDL_STRCO	P42534 streptomyc
8	172.5	18.0	563	2 Q6J6N0_9ARAC	O6J6N0 araneus ven
9	169	17.6	617	2 Q5B4Q0_EMENI	O5B4Q0 aspergillus
10	168.5	17.6	1002	2 Q9BIU8_ARGTR	Q9BIU8 argiope tri
11	167.5	17.5	922	2 Q43367_MYTGD	O43367 mytilus edu
12	165.5	17.3	922	2 Q8MW53_MYTGA	Q8MW53 mytilus gal
13	162	16.9	420	2 Q6RF45_ARATH	O6RF45 arabidopsis
14	162	16.9	651	2 Q9BIU9_ARGTR	Q9BIU9 argiope tri
15	161	16.8	524	2 Q8I7U1_9ARAC	O8I7U1 araneus ven
16	160.5	16.7	813	2 Q636W4_BACCC	O636W4 bacillus ce
17	160.5	16.7	1168	2 Q6HF99_BACHK	Q6HF99 bacillus th
18	160.5	16.7	1408	2 Q7U022_MYPCPL	Q7U022 mycobacteri
19	160.5	16.7	2249	2 Q9NHW4_NEPCL	Q9NHW4 nephila cla
20	160	16.7	604	2 Q9L252_STRCO	Q9L252 streptomyc
21	160	16.7	871	2 Q43358_NEPCL	O43358 nephila cla
22	159.5	16.6	1884	2 Q9NHW2_9ARAC	Q9NHW2 nephila ina
23	159	16.6	904	2 Q76271_MYTGD	O76271 mytilus edu
24	158.5	16.5	678	2 Q93486_ONCMY	Q93486 oncorhynch
25	158.5	16.5	1458	2 Q910B9_ONCMY	Q910B9 oncorhynch
26	157.5	16.4	597	2 Q58MV1_9CAUD	O58MV1 cyanophage
27	157	16.4	410	2 Q16988_ARADI	Q16988 araneus dia
28	157	16.4	626	2 Q9NHW1_9ARAC	Q9NHW1 nephila ina
29	156.5	16.3	1046	1 IF2_STRAW	Q82K53 streptomyc
30	155.5	16.2	145	2 Q61351_DROME	O61351 drosophila
31	155.5	16.2	641	1 EBN1_EBV	P03211 Epstein-Bar

RESULT 1					ALIGNMENTS				
ID	CTG1B_HUMAN	STANDARD;	PRT;	180 AA.	Q777E1	human herpe	Q777E1	human herpe	
AC	P78358;				Q8MW54	mytilus gal	Q8MW54	mytilus gal	
DT	15-DEC-1998 (Rel. 37, Created)				O16161	MYTED	O16161	MYTED	
DT	15-DEC-1998 (Rel. 37, Last sequence update)				Q8LWX2	BACAN	Q8LWX2	BACAN	
DT	10-MAY-2005 (Rel. 47, Last annotation update)				Q6HV27	BACAN	Q6HV27	BACAN	
DE	Cancer/testis antigen 1B (Autoimmunogenic cancer/testis antigen NY-ESO-1).				Q8MW55	mytilus gal	Q8MW55	mytilus gal	
GN	Name=CTAG1B; Synonyms=CTAG, CTAG1;				Q9BIT7	MYCTU	Q9BIT7	MYCTU	
OS	Homo sapiens (Human).				Q6NZ15	BRARE	Q6NZ15	BRARE	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				Q6PEI9	BRARE	Q6PEI9	BRARE	
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Hominidae;				Q6IVJ4	CTOIN	Q6IVJ4	CTOIN	
OC	Homo.				734		734		
OX	NCBI_TaxID=9606;				15.9		15.9		
RN	[1]				779		779		
RP	NCBI_TaxID=9606;				1938		1938		
RX	MEDLINE=97203161; PubMed=9050879; DOI=10.1073/pnas.94.5.1914;				219		219		
RA	Chen Y.-T., Scanlan M.J., Sahin U., Tuerci O., Gure A.O., Tsang S., Williamson B., Stockert E., Pfeundscher M., Old L.J.;								
RT	"A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening."								
RL	Proc. Natl. Acad. Sci. U.S.A. 94:1914-1918(1997).								
RN	[2]								
RP	NUCLEOTIDE SEQUENCE.								
RC	TISSUE=Melanoma;								
RX	MEDLINE=98289662; PubMed=9626360;								
RA	Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A., de Plaen E., Boon T.;								
RT	"LAGE-1, a new gene with tumor specificity."								
RL	Int. J. Cancer 76:903-908(1998).								
RN	[3]								
RP	NUCLEOTIDE SEQUENCE.								
RX	MEDLINE=98430882; PubMed=9759882;								
RA	Wang R.-F., Johnston S.L., Zeng G., Topalian S.L., Schwartzentruber D.J., Rosenberg S.A.;								
RT	"A breast and melanoma-shared tumor antigen: T cell responses to antigenic peptides translated from different open reading frames."								
RL	J. Immunol 161:3596-3606(1998).								
CC	-!- TISSUE SPECIFICITY: Expressed in testis and ovary and in a wide variety of cancers. Detected in uterine myometrium.								
CC	-!- SIMILARITY: Belongs to the CTAG family.								
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.								
DR	EMBL; U87459; AAB49693.1; -; mRNA.								
DR	EMBL; AJ003149; CA05908.1; -; mRNA.								
DR	EMBL; AF038567; AAD05202.1; -; mRNA.								
DR	FGNC; HGNC:2491; CTAG1B.								
DR	MIM; 300156; -.								

ID	Q9NY13_HUMAN	PRELIMINARY;	PRT;	142	AA.
AC	Q9NY13;				
AD	01-OCT-2000	(TrEMBLrel. 15, Created)			
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)			
DE	Hypothetical protein LAGE-2	(Fragment).			
GN	Names=LAGE-2;				
OS	Homo sapiens	(Human).			
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria;	Euarchontoglires; Primates; Catarrhini; Hominidae			
OC	Homo.				
OX	NCBI_TaxID=9606;				
LN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RP	Lethe B.G.;				
RL	Submitted (MAR-2000)	to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ275978;	CAB76945.1; -; mRNA.			
KW	Hypothetical protein.				
FT	NON_TER				
FT	1				
SQ	SEQUENCE	142 AA; 13895 MW; 27EBE922AC4ACC7B	CRC64;		
	Query Match	60.6%;	Score 581;	DB 2;	Length 142;
	Best Local Similarity	94.8%;	Pred. No. 2,5e-34;		
	Matches 109;	Conservative	2;	Mismatches	4; Indels 0; Gaps
Qy	27	PGNAGGPGCEAGATGGRGPRGAGAAARASGPGGAGPRGPHGGAASGLNGCCRCGARGPESR			
Db	1	PGNAGGPGCEAGATGGRGPRGAGAAARASGPGGAGPRGPHGGAASGLNGCCRCGARGPESR			
Qy	87	LLEFYLMFPATPMAELARRSLAQADAPLPVPGVLLKEFTVSGNLTIRLTAAAD 141			
Db	61	LLEFYLMFPATPMAELARRSLAQADAPLPVPGVLLKEFTVSGNLTMSVQDDQ 115			
RESULT 5					
ID	ITBA2_HUMAN	STANDARD;	PRT;	143	AA.
AC	Q14657; Q81Z78;				
AD	01-NOV-1997	(Rel. 35, Created)			
DT	25-OCT-2004	(Rel. 45, Last sequence update)			
DT	10-MAY-2005	(Rel. 47, Last annotation update)			
DE	ITBA2 protein (ESO-3 protein)	(DXS9879E).			
GN	Names=ITBA2; Synonyms=ESO3;				
OS	Homo sapiens	(Human).			
OC	Eukaryota; Metazoa;	Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria;	Euarchontoglires; Primates; Catarrhini; Hominidae			
OC	Homo.				
OX	NCBI_TaxID=9606;				
LN	[1]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].				
RP	TISSUE=Blood, and Pancreas;				
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Caavaant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Burttfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,				
RA	Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	"generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RN	[Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002)].				
RN	[2]				

```
RE NUCLEOTIDE SEQUENCE OF 2-143, AND TISSUE SPECIFICITY.
RC TISSUE=Liver;
RX MEDLINE=96374823; PubMed=8786111; DOI=10.1006/geno.1996.0293;
RA Faranda S., Frattini A., Zucchi I., Patrosso C., Milanesi L.,
RA Montagna C., Vezzoni P.;
RT "Characterization and fine localization of two new genes in Xq28 using
RT the genomic sequence/EST database screening approach.";
RL Genomics 34:323-327(1996).
RN [3]
RP IDENTIFICATION, AND TISSUE SPECIFICITY.
RX PubMed=12384295; DOI=10.1016/S0378-1119(02)00879-X;
RA Alpen B., Guere A.O., Scanlan M.J., Old L.J., Chen Y.-T.;
RT "A new member of the NY-ESO-1 gene family is ubiquitously expressed in
RT somatic tissues and evolutionarily conserved.";
RL Gene 297:141-149(2002).
CC -I- TISSUE SPECIFICITY: Ubiquitous.
CC -I- SIMILARITY: Belongs to the CTAG family.
CC -I- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 54.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR ENBL; BC015744; AAH15744.2; -; mRNA.
DR ENBL; BC062330; AAH62330.1; -; mRNA.
DR ENBL; X92896; CAAG3489.1; ALT_FRAME; mRNA.
DR Ensembl; ENSG00000196976; Homo sapiens.
DR MTM; 300060; -.
SQ SEQUENCE 143 AA; 14804 MW; AD164559371449F8 CRC64;

Query Match 23.1%; Score 222; DB 1; Length 143;
Best Local Similarity 42.1%; Pred. No. 1.6e-08;
Matches 67; Conservative 14; Mismatches 52; Indels 26; Gaps 6;

QY 14 DADGPGGPGIPDGPNGAGGCEAGATGGRGPRGAGAAASGPGGAP---RGPHGGAA 69
Db 3 DADADAG-----GGADGGDGR-----GGHSCRG-GVDTAAAPAGGAPPAHAPGPGRDAA 50

QY 70 SGLNGCCRCGARGPESRLLEFYLLAMPFPATPMEAEIARRSLAQDAPPLPVPGVLLKEFTVS 129
Db 51 S-----AARGSRMRPHFTLSVFPPTPLEAEIAGSLAPDAE--PHQRVVVGKDLTVS 100

QY 130 GNILTRLTADHRLQQLSISSCLOQLSLMWITQCFLP 168
Db 101 GRILVVRWKAEDCRLRLISVINFLDQLSLVVRTMQRFGP 139

RESULT 7
HYDL_STRCO
ID HYDL_STRCO STANDARD; PRT; 627 AA.
AC P42534; Q9S2L7;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Putative polyketide hydroxylase (EC 1.14.13.-) (White ORF VIII).
GN OrderedLocusNames=SC05321; ORFNames=SC6G9.12c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kießer H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble T., Larke L., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kießer T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajadream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-255.
RC STRAIN=A3(2);
RX MEDLINE=94075247; PubMed=8253693;
RA Blanco G., Pereda A., Brian P., Mendez C., Chater K.F., Salas J.A.;
RT "A hydroxylase-like gene product contributes to synthesis of a
RT polyketide spore pigment in Streptomyces halstedii.";
RL J. Bacteriol. 175:8043-8048(1993).
CC -I- FUNCTION: Involved in developmentally regulated synthesis of a
CC compound biosynthetically related to polyketide antibiotics which
CC is essential for spore color in Streptococcus coelicolor.
CC -I- COFACTOR: FAD (By similarity).
CC -I- SIMILARITY: Belongs to the pheA/tfdB FAD monooxygenase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR ENBL; AL939123; CAB45603.1; -; Genomic_DNA.
DR ENBL; X74213; CAAS2289.1; -; Genomic_DNA.
DR PIR; T35608; T35608.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002938; mOase FAD bd.
DR InterPro; IPR003042; Rng hydrolase.
DR InterPro; IPR012336; Thioresoxin-like.
DR Pfam; PF01494; FAD_binding_3; 1.
DR PRINTS; PR00420; RNMNOXGNASE.
QY 14 DADGPGGPGIPDGPNGAGGCEAGATGGRGPRGAGAAASGPGGAP---RGPHGGAA 69
Db 3 DADADAG-----GGADGGDGR-----GGHSCRG-GVDTAAAPAGGAPPAHAPGPGRDAA 50

QY 70 SGLNGCCRCGARGPESRLLEFYLLAMPFPATPMEAEIARRSLAQDAPPLPVPGVLLKEFTVS 129
Db 51 S-----AARGSRMRPHFTLSVFPPTPLEAEIAGSLAPDAE--PHQRVVVGKDLTVS 100

QY 130 GNILTRLTADHRLQQLSISSCLOQLSLMWITQCFLP 168
Db 101 GRILVVRWKAEDCRLRLISVINFLDQLSLVVRTMQRFGP 139

RESULT 6
QSHY39_HUMAN
ID QSHY39_HUMAN PRELIMINARY; PRT; 143 AA.
AC QSHY39;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE L antigen family, member 3.
GN Name=LAGE3; ORFNames=XX-FW81657B9.4-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Heath P.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR ENBL; BX936365; CAI43195.1; -; Genomic DNA.
SQ SEQUENCE 143 AA; 14804 MW; AD164559371449F8 CRC64;

Query Match 23.1%; Score 222; DB 2; Length 143;
Best Local Similarity 42.1%; Pred. No. 1.6e-08;
```


OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilinae; Mytilus.
 OK NCBI_TaxID=29158;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22038007; PubMed=12042339;
 RA Lucas J.M., Vaccaro E., Waite J.H.;
 RT "A molecular, morphometric and mechanical comparison of the structural
 RT elements of byssus from Mytilus edulis and Mytilus
 RT galloprovincialis."
 RL J. Exp. Biol. 205:1807-1817(2002).
 DR EMBL: AF448526; AAM34601.1; -; mRNA.
 DR GO: GO:0005737; C:cytoplasm; IEA.
 DR GO: GO:0006817; P:phosphate transport; IEA.
 DR InterPro: IPR008160; Collagen.
 DR Pfam: PF01391; Collagen; 6.
 KW Collagen.
 SQ SEQUENCE 922 AA; 80750 MW; CA281AE0975BB7D4 CRC64;

 Query Match 17.3%; Score 165.5; DB 2; Length 922;
 Best Local Similarity 47.6%; Pred. No. 0.00099;
 Matches 39; Conservative 1; Mismatches 29; Indels 13; Gaps 3;

 QY 5 GRGTGSGTGDADGPGGPIPDGPGNA--GGPGAGATGCGPRGAGAAASGPGGAPR 62
 DB 174 GGTGG-EGQDPGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG 232

 QY 63 GPHGGAASGLNGCCRCGARGPE 84
 DB 233 GPPGPA-----GPGGPE 244

 RESULT 13
 Q6RF45 ARATH
 ID Q6RF45 ARATH PRELIMINARY; PRT; 420 AA.
 AC Q6RF45;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=At2g04170;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OK NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Xiao Y., Smith S.R., Ishmael N., Ayele M., Kumar N., Redman J.,
 RA Riedmuller S., Utterback T., Whitelaw C.A., Fraser C.M., Town C.D.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Underwood B.A., Xiao Y., Moskal W., Torian U., Redman J., Wu H.C.,
 RA Utterback T., Town C.D.;
 RT "Arabidopsis thaliana ORF clones of hypothetical genes."
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY501356; AAR99368.1; -; mRNA.
 DR EMBL: AY649317; AAT69234.1; -; mRNA.
 DR InterPro: IPR002083; MATH.
 DR Pfam: PF00917; MATH; 2.
 DR SMART: SM00061; MATH; 2.
 DR PROSITE: PS0144; MATH; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 420 AA; 44404 MW; 6B108942B389EEA9 CRC64;

 Query Match 16.9%; Score 162; DB 2; Length 420;
 Best Local Similarity 50.0%; Pred. No. 0.00086;
 Matches 42; Conservative 0; Mismatches 32; Indels 10; Gaps 5;

 QY 5 GRGTG---GSTGDADGPGGPI--PDGPGNAGGPGAGATGCGPRGAG--AARASGPGG 59
 DB 19 GRGGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG 74

QY	60	APRPHGGAASGLNCCRCRGARGP	83
DB	75	-PRGPRPGGGGPGPGWSPGRGP	97

RESULT 14

Q9BIU9_ARGTR	PRELIMINARY;	PRT;	651 AA.
ID Q9BIU9;			
AC Q9BIU9;			
DT 01-JUN-2001	(TREMBLrel. 17, Created)		
DT 01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT 01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	Flagelliform silk protein (Fragment).		
GN	Name=Flag;		
OS	Argiope trifasciata (Banded garden spider).		
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;		
OX	Araneomorphae; Entelegynae; Araneoidea; Araneidae; Argiope.		
NCBI_TaxID=156845;			
RN [1]			
RN NUCLEOTIDE SEQUENCE.			
RX MEDLINE=21179804; PubMed=11283372; DOI=10.1126/science.1057561;			
RA Gatesy J., Hayashi C., Motriuk D., Woods J., Lewis R.;			
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AF350264; AKK30593.1; -; mRNA.			
FT NON TER	1		
SQ SEQUENCE	651 AA; 52360 MW; 77CEECE8757DE48DA CRC64;		
	Query Match	16.9%;	Score 162; DB 2; Length 651;
	Best Local Similarity	44.0%;	Pred: No. 0.0013;
	Matches	37; Conservative	2; Mismatches 37; Indels 8; Gaps 1;

QY	7	GTGGSTGDADGGGGIGIPDGGGNAGFGGEAGATCGRPRGAGAARASGGGGAP----	61
DB	168	GPGGGAGAGPGGPGGGGGLGGVGGAGTGGVGPGAGGVGPGSGAGTGTPVSVS	227
QY	62	---RGPHGGAASGLNCCRCRGARG	82
		:	
DB	228	TVSVGGAGGPGAGGPGAGGAGG	251
		:	

RESULT 15

Q8I7UI_9ARAC	PRELIMINARY;	PRT;	524 AA.
ID Q8I7UI_9ARAC			
AC Q8I7UI;			
DT 01-MAR-2003	(TREMBLrel. 23, Created)		
DT 01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT 01-JUN-2003	(TREMBLrel. 24, Last annotation update)		
DE	Major ampullate gland dragline silk protein-1 (fragment).		
OS	Araeneus ventricosus.		
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;		
OX	Araneomorphae; Entelegynae; Araneoidea; Araneidae; Araeneus.		
NCBI_TaxID=182803;			
RN [1]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Liu Z.-S., Ren H.-L., Pan F.-G.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN [2]			
RN NUCLEOTIDE SEQUENCE.			
RA Ren H.-L., Liu Z.-S., Pan F.-G., Zhao J., Li Y.-S.;			
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AY174110; AAN85280.1; -; mRNA.			
DR GO; GO:0005737; C:cycloplasm; IEA.			
DR GO; GO:0006817; P:phosphate transport; IEA.			
DR InterPro; IPR008160; Collagen.			
RN			

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:17:37 ; Search time 50 Seconds
(without alignments)
297.632 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAERGTTGGTGDADPGG.....WITQCFVFLAQPSPGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5 COMB.pap.*
 - 2: /cgn2_6/prodata/1/iaa/6 COMB.pap.*
 - 3: /cgn2_6/prodata/1/iaa/H COMB.pap.*
 - 4: /cgn2_6/prodata/1/iaa/PCBUS COMB.pap.*
 - 5: /cgn2_6/prodata/1/iaa/RE COMB.pap.*
 - 6: /cgn2_6/prodata/1/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	180	1	US-08-791-495-9
2	959	100.0	180	2	US-09-392-714-25
3	959	100.0	180	2	US-09-165-546D-15
4	959	100.0	180	2	US-09-341-829A-9
5	959	100.0	180	2	US-09-849-602-30
6	953	99.4	180	2	US-08-937-263B-8
7	953	99.4	180	2	US-09-751-798-8
8	814	84.9	180	1	US-08-791-495-7
9	814	84.9	180	2	US-09-341-829A-7
10	589	61.4	210	1	US-08-791-495-5
11	589	61.4	210	2	US-09-341-829A-5
12	160	16.7	745	1	US-09-010-928B-28
13	160	16.7	870	1	US-09-010-928B-2
14	156.5	16.3	633	1	US-08-642-255-73
15	156.5	16.3	1065	1	US-08-642-255-72
16	155.5	16.2	235	1	US-08-529-190B-1
17	155.5	16.2	641	2	US-09-249-585A-3
18	155.5	16.2	641	2	US-09-410-399-4
19	151.5	15.8	1057	2	US-08-931-820-1
20	151.5	15.8	1057	2	US-10-153-469A-16
21	151.5	15.8	1057	2	US-10-153-469A-20
22	151.5	15.8	1057	2	US-10-104-889-16
23	151.5	15.8	1057	2	US-10-104-889-20
24	151.5	15.8	1107	2	US-10-153-469A-11
25	151.5	15.8	1107	2	US-10-104-889-11
26	151.5	15.8	1169	2	US-10-153-469A-6
27	151.5	15.8	1169	2	US-10-104-889-6

28	151.5	15.8	1171	2	US-10-153-469A-8	Sequence 8, Appli
29	151.5	15.8	1171	2	US-10-104-889-8	Sequence 8, Appli
30	151.5	15.8	1341	2	US-08-963-825-18	Sequence 18, Appli
31	151.5	15.8	1341	2	US-09-500-811-18	Sequence 18, Appli
32	151.5	15.8	1341	2	US-09-570-573-18	Sequence 18, Appli
33	151.5	15.8	1341	2	US-09-548-608-18	Sequence 18, Appli
34	151.5	15.8	1388	2	US-10-153-469A-10	Sequence 10, Appli
35	151.5	15.8	1388	2	US-10-104-889-10	Sequence 10, Appli
36	151.5	15.8	1461	2	US-09-585-887-9	Sequence 9, Appli
37	151.5	15.8	1461	2	US-09-289-578-9	Sequence 9, Appli
38	151.5	15.8	1464	2	US-09-331-347C-21	Sequence 21, Appli
39	150.5	15.7	907	1	US-09-010-928B-4	Sequence 4, Appli
40	150	15.6	1008	2	US-09-219-849-8	Sequence 8, Appli
41	150	15.6	1065	1	US-08-642-255-80	Sequence 80, Appli
42	150	15.6	1065	2	US-08-642-246-16	Sequence 16, Appli
43	150	15.6	1065	2	US-09-451-206-16	Sequence 16, Appli
44	150	15.6	1065	4	PCT-US96-06229-16	Sequence 16, Appli
45	149	15.5	166	2	US-09-297-269-41	Sequence 41, Appli

ALIGNMENTS

RESULT 1

US-08-791-495-9

; Sequence 9, Application US/08791495

; Patent No. 5811519

; GENERAL INFORMATION:

; APPLICANT: Leth, Bernard

; APPLICANT: Lucas, Sophie

; APPLICANT: De Smet, Charles

; APPLICANT: Godelaine, Daniele

; APPLICANT: Boon-Falleur, Thierry

; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

; STREET: 600 Atlantic Avenue

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/791,495

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Van Amsterdam, John R.

; REGISTRATION NUMBER: 40,212

; REFERENCE/DOCKET NUMBER: L0461/7005

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-720-3500

; TELEFAX: 617-720-2441

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 180 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-791-495-9

Query Match 100.0%; Score 959; DB 1; Length 180;

Best Local Similarity 100.0%; Pred. No. 9.7e-79;

Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAERGTTGGTGDADPGGCPIDPGGNAGGPBAGATGGRGPRGAGARASGPGGA 60

DB 1 MQAERGTTGGTGDADPGGCPIDPGGNAGGPBAGATGGRGPRGAGARASGPGGA 60

QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
DE 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180

RESULT 2

US-09-392-714-25
; Sequence 25, Application US/09392714A
; Patent No. 6886147
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Gure, Ali O.
; APPLICANT: Williamson, Barbara
; APPLICANT: Chen, Yao-Taeng
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Cancer Associated Antigens and Uses
; FILE REFERENCE: L0461/7062
; CURRENT APPLICATION NUMBER: US/09/392,714A
; EARLIER FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: PCT/US98/14679
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-392-714-25

Query Match 100.0%; Score 959; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 9.7e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60
DB 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
DB 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180

RESULT 3

US-09-165-546D-15
; Sequence 15, Application US/09165546D
; Patent No. 6723832
; GENERAL INFORMATION:
; APPLICANT: Knuch, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
; Gure, Ali, Old, Lloyd, Ritter, Gerd
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CL
; USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FULBRIGHT & JAWORSKI LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10158
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM

; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,546D
; FILING DATE: 02-Oct-1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6723832man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15
US-09-165-546D-15

Query Match 100.0%; Score 959; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 9.7e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60
DB 1 MQAEGRTGGSTGDADGPGGPGIPDGPNGAGPGEAGATGGRPGAGAAASGPGGGA 60
QY 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
DB 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEAEALARRSLAQAADAPPLVPVG 120
QY 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180
DB 121 VLLKFTVSGNLTIRLTAAADHRQLQSLSSCQLQSLMLWITQCFLPVFLAQPSPGQRR 180

RESULT 4

US-09-341-829A-9
; Sequence 9, Application US/09341829A
; Patent No. 6794131
; GENERAL INFORMATION:
; APPLICANT: Leth, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LAGE-1 TUMOR ASSOCIATED NUCLEIC ACIDS
; FILE REFERENCE: L0461/7066
; CURRENT APPLICATION NUMBER: US/09/341,829A
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 08/791,495
; PRIOR APPLICATION NUMBER: PCT/US98/01445
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-341-829A-9

Query Match 100.0%; Score 959; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 9.7e-79;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MQAEGRTGSGTDDADGPGGPGIGDGGCGNAGGPGCEAGATGGRGPRGAGAAARASGPGGGA	60
Db	1	MQAEGRTGSGTDDADGPGGPGIGDGGCGNAGGPGCEAGATGGRGPRGAGAAARASGPGGGA	60
Qy	61	PRGPHGGAAGLNGCCRCGARGPESRLLEFYLAHPFATPMEAEIARRSLAQDAPPLVPVG	120
Db	61	PRGPHGGAAGLNGCCRCGARGPESRLLEFYLAHPFATPMEAEIARRSLAQDAPPLVPVG	120
Qy	121	VLLKETVSGNLTITRLTAADHQRLQLSISSCLQLSLMWITOCFLPVFLAOPPSGRR	180
Db	121	VLLKETVSGNLTITRLTAADHQRLQLSISSCLQLSLMWITOCFLPVFLAOPPSGRR	180

RESULT 5

```

US-09-849-602-30
; Sequence 30, Application US/09849602
; Patent No. 6794501
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew J.
; APPLICANT: Old, Lloyd J.
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Chen, Yao-Tseng
; TITLE OF INVENTION: Colon Cancer Antigen Panel
; FILE REFERENCE: L0461/7105(JRV)
; CURRENT APPLICATION NUMBER: US/09/849,602
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-602-30

```

Query Match	100.0%	Score 959;	DB 2;	Length 180;
Best Local Similarity	100.0%	Pred. No. 9.7e-79;		
Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 6

US-08-937-263B-8
Sequence 8, Application US/08937263B
Patent No. 6274145
GENERAL INFORMATION:
APPLICANT: Chen, Yao-Tseung; Scanlan, Matthew;
APPLICANT: Gure, Ali; Old, Lloyd J.; Jager, Elke;
APPLICANT: Alexander, Knuth; Drijfhout, Jan W.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE
TITLE OF INVENTION: ENCODING CANCER ASSOCIATED ANTIGEN, THE ANTIGEN
TITLE OF INVENTION: ITSELF, AND USES THEREOF
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM

```

? OPERATING SYSTEM: PC-DOS
? SOFTWARE: WordPerfect
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/937,263B
? FILING DATE: September 15, 1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/725,182
? FILING DATE: October 3, 1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Sinn, Eric, Patent Agent
? REGISTRATION NUMBER: 40,177
? REFERENCE/DOCKET NUMBER: LUD 5466.1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 318-3000
? TELEFAX: (212) 752-5958
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 180
? TYPE: amino acid
? TOPOLOGY: linear
?
US-08-937-263B-8

Query Match          99.4%; Score 953; DB 2; Length 180;
Best Local Similarity 99.4%; Pred. No. 3.4e-76;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1    MQAEGRTGGSTGDADGFGGFGIGIDGPGGNAGGPGEAGATGGRGPRGAGAARASGPGGGA 60
        |||
Db      1    MQAEGRTGGSTGDADGFGGFGIGIDGPGGNAGGPGEAGATGGRAPRGAGAARASGPGGGA 60
        |||

Qy      61   PRGFPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMPEAEIARSLAQDAAPLPVPVG 120
        |||
Db      61   PRGFPHGGAASGLNGCCRCGARGPESRLLEFYLAMPFATPMPEAEIARSLAQDAAPLPVPVG 120
        |||

Qy      121  VLKKEFTVSGNIIIRLTAAHDHQLQLSISSCIQQLSLLMWITOCFLPVFLAQPPSGQR 180
        |||
Db      121  VLKKEFTVSGNIIIRLTAAHDHQLQLSISSCIQQLSLLMWITOCFLPVFLAQPPSGQR 180
        |||

```

RESULT 7

US-09-751-798-8
; Sequence 8, Application US/09751798
; Patent No. 6525177
; GENERAL INFORMATION:
; APPLICANT: Stockert, Elisabeth; Jager, Elke;
; APPLICANT: Chen, Yao-tseng; Scanlan, Matthew;
; APPLICANT: Knuth, Alexander; Old, Lloyd J.
; TITLE OF INVENTION: Antibodies Which Bind to NY-ESO-1 Cancer
; TITLE OF INVENTION: Associated Proteins, Uses Thereof,
; TITLE OF INVENTION: Truncated Forms of NY-ESO-1, and HLA
; TITLE OF INVENTION: Binding Peptides Derived Therefrom
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P.
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09751,798
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/062,422
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/725,182
; FILING DATE: October 3, 1996


```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 652517man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5466.3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3168
; TELEFAX: (212) 752-5958
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180
; TYPE: amino acid
; TOPOLOGY: linear
;
US-09-751-798-8

Query Match          99.4%; Score 953; DB 2; Length 180;
Best Local Similarity 99.4%; Pred.No.3.4e-78;
Matches 179; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 MQAEGRTGSGTGDADPGGPGIGIPDGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
      |||
Db      1 MQAEGRTGSGTGDADPGGPGIGIPDGGNAGGPGGAGATGGRAPRGAGAAASGPGGGA 60
      |||

QY      61 PRGPHGGAASGLNGCCRCGARGPESRLLFFYLAMPFATPMEAEIARSLAODAPLPVPG 120
      |||
Db      61 PRGPHGGAASGLNGCCRCGARGPESRLLFFYLAMPFATPMEAEIARSLAODAPLPVPG 120
      |||

CY      121 VLLKEFTVSGNLTIRLTADHROLOLSISCLQOLSLLMWITOCFLPVFLAQPPSGQRR 180
      |||
Db      121 VLLKEFTVSGNLTIRLTADHROLOLSISCLQOLSLLMWITOCFLPVFLAQPPSGQRR 180
      |||

RESULT 8
US-08-791-495-7
; Sequence 7, Application US/08791495
; Patent No. 5811519
; GENERAL INFORMATION:
; APPLICANT: Leth , Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boor-Falleur, Thierry
; TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,495
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: L0461/7005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-791-495-7

```

```
/ STREET: 600 Atlantic Avenue
/ CITY: Boston
/ STATE: MA
/ COUNTRY: USA
/ ZIP: 02210
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/791,495
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Van Amsterdam, John R.
/ REGISTRATION NUMBER: 40,212
/ REFERENCE/DOCKET NUMBER: L0461/7005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-720-3500
/ TELEFAX: 617-720-2441
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 210 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-08-791-495-5
/
/ Query Match 61.4%; Score 589; DB 1; Length 210;
/ Best Local Similarity 76.6%; Pred. No. 2e-45;
/ Matches 108; Conservative 15; Mismatches 18; Indels 0; Gaps 0;
/
/ QY 1 MQAEGRTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/ DB 1 MQAEGGCTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/
/ QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/ DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/
/ QY 121 VLLKEFTVSGNLTIRLTAAD 141
/ DB 121 AVLKDFTVSGNLLFMSVRDQD 141
/
/ RESULT 12
/ US-09-010-928B-28
/ Sequence 28, Application US/09010928B
/ Patent No. 5994099
/ GENERAL INFORMATION:
/ APPLICANT: Lewis, Randolph V
/ APPLICANT: Hayaashi, Cheryl Y
/ TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
/ TITLE OF INVENTION: CODING THEREFOR
/ NUMBER OF SEQUENCES: 29
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
/ STREET: 8110 GATEHOUSE RD. SUITE 500E
/ CITY: FALLS CHURCH
/ STATE: VIRGINIA
/ COUNTRY: UNITED STATES OF AMERICA
/ ZIP: 22042
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/010,928B
/ FILING DATE: 22-JAN-1998
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Murphy Jr., Gerald M
/ REGISTRATION NUMBER: 28977
/ REFERENCE/DOCKET NUMBER: 1447-109P
/ INFORMATION FOR SEQ ID NO: 28:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 745 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/
/ US-09-010-928B-28
/
/ Query Match 16.7%; Score 160; DB 1; Length 745;
/ Best Local Similarity 52.4%; Pred. No. 2.9e-06;
/ Matches 43; Conservative 1; Mismatches 32; Indels 6; Gaps 5;
/
/ QY 5 GRGTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGG-GAPR 62
/ DB 508 GFGGFGGPGGFGGPGGAGGPGYGG-AGGPGGAGGPGYGGAGG---PYGPGGAGGPGY 564
/
/ QY 63 GPHG-GAASGLNGCCRCGARGP 83
/ DB 565 GFGGAGGSGYLGAGGSGGVGP 586
/
/ RESULT 13
/ US-09-010-928B-2
/ Sequence 2, Application US/09010928B
/ Patent No. 5994099
/ GENERAL INFORMATION:
/ APPLICANT: Lewis, Randolph V
/ APPLICANT: Hayaashi, Cheryl Y
/ TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
/
/ QY 1 MQAEGRTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/ DB 1 MQAEGGCTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/
/ QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/ DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/
/ QY 121 VLLKEFTVSGNLTIRLTAAD 141
/ DB 121 AVLKDFTVSGNLLFMSVRDQD 141
/
/ RESULT 11
/ US-09-341-829A-5
/ Sequence 5, Application US/09341829A
/ Patent No. 6794131
/ GENERAL INFORMATION:
/ APPLICANT: Leth,, Bernard
/ APPLICANT: Lucas, Sophie
/ APPLICANT: De Smet, Charles
/ APPLICANT: Godelaine, Daniele
/ APPLICANT: Boon-Falleur, Thierry
/ TITLE OF INVENTION: LAGR-1 TUMOR ASSOCIATED NUCLEIC ACIDS
/ FILE REFERENCE: L0461/7066
/ CURRENT APPLICATION NUMBER: US/09/341,829A
/ CURRENT FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: US 08/791,495
/ PRIOR APPLICATION NUMBER: PCT/US98/01445
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5
/ LENGTH: 210
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
/ US-09-341-829A-5
/
/ Query Match 61.4%; Score 589; DB 2; Length 210;
/ Best Local Similarity 76.6%; Pred. No. 2e-45;
/ Matches 108; Conservative 15; Mismatches 18; Indels 0; Gaps 0;
/
/ QY 1 MQAEGRTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/ DB 1 MQAEGGCTGSGTDADPGGPGIPDGPNGAGGPGGAGATGCGPRGAGAAASGPGGA 60
/
/ QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/ DB 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPAMPFATPMEAEALARRSLAQAADAPLPVPG 120
/
/ QY 121 VLLKEFTVSGNLTIRLTAAD 141
/ DB 121 AVLKDFTVSGNLLFMSVRDQD 141
/
/ RESULT 11
/ US-09-341-829A-5
/ Sequence 5, Application US/09341829A
/ Patent No. 6794131
/ GENERAL INFORMATION:
/ APPLICANT: Leth,, Bernard
/ APPLICANT: Lucas, Sophie
/ APPLICANT: De Smet, Charles
/ APPLICANT: Godelaine, Daniele
/ APPLICANT: Boon-Falleur, Thierry
/ TITLE OF INVENTION: LAGR-1 TUMOR ASSOCIATED NUCLEIC ACIDS
/ FILE REFERENCE: L0461/7066
/ CURRENT APPLICATION NUMBER: US/09/341,829A
/ CURRENT FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: US 08/791,495
/ PRIOR APPLICATION NUMBER: PCT/US98/01445
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5
/ LENGTH: 210
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/
/ US-09-341-829A-5
/
/ Query Match 61.4%; Score 589; DB 2; Length 210;
/ Best Local Similarity 76.6%; Pred. No. 2e-45;
/ Matches 108; Conservative 15; Mismatches 18; Indels 0; Gaps 0;
```


This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:29:28 ; Search time 174 Seconds
(without alignments)
432.237 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAEGRTGGTGDADPGG.....WITQCFLPVFLAQPSPGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA_Main:

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pap.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pap.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pap.*
- 4: /cgn2_6/prodata/1/pubpaa/US10_PUBCOMB.pap.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pap.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	959	100.0	180	3	US-09-849-602-30
2	959	100.0	180	4	US-10-207-655-71
3	959	100.0	180	4	US-10-026-066-3
4	959	100.0	180	4	US-10-117-937-74
5	959	100.0	180	4	US-10-295-027-386
6	959	100.0	180	4	US-10-296-734-832
7	959	100.0	180	4	US-10-188-832-139
8	959	100.0	180	4	US-10-777-053-11
9	959	100.0	180	4	US-10-751-088-15
10	959	100.0	180	4	US-10-657-022-74
11	959	100.0	180	4	US-10-837-217-11
12	959	100.0	180	5	US-10-877-373-9
13	959	100.0	180	5	US-10-723-860-1270
14	959	100.0	180	5	US-10-871-708-7
15	959	100.0	180	5	US-10-895-523-3
16	959	100.0	180	5	US-10-182-506A-3
17	959	100.0	180	5	US-10-756-149-5024
18	959	100.0	180	6	US-11-067-064-74
19	959	100.0	180	6	US-11-067-159-74
20	959	100.0	337	3	US-09-821-883-27
21	959	100.0	337	6	US-11-144-912-27
22	954	99.5	179	5	US-10-482-029-202
23	953	99.4	180	3	US-09-751-798-8
24	953	99.4	180	4	US-10-023-182-8
25	918	95.7	180	4	US-10-364-614-14
26	821	85.6	180	4	US-10-117-937-75
27	821	85.6	180	4	US-10-657-022-75

28	821	85.6	180	6	US-11-067-064-75	Sequence 75, Appl
29	821	85.6	180	6	US-11-067-159-75	Sequence 75, Appl
30	814	84.9	180	4	US-10-146-473-69	Sequence 69, Appl
31	814	84.9	180	4	US-10-296-734-834	Sequence 834, App
32	814	84.9	180	4	US-10-468-406-4	Sequence 4, Appl
33	814	84.9	180	5	US-10-877-373-7	Sequence 7, Appl
34	631.5	65.8	135	4	US-10-295-027-388	Sequence 141, App
35	631.5	65.8	135	4	US-10-188-832-141	Sequence 388, App
36	596	62.1	210	4	US-10-157-031-88	Sequence 88, Appl
37	595	62.0	210	4	US-10-117-937-76	Sequence 76, Appl
38	595	62.0	210	4	US-10-657-022-76	Sequence 76, Appl
39	595	62.0	210	6	US-11-067-159-76	Sequence 76, Appl
40	595	62.0	210	6	US-11-067-159-76	Sequence 5, Appl
41	589	61.4	210	5	US-10-877-373-5	Sequence 17, Appl
42	532	55.5	123	4	US-10-777-053-17	Sequence 17, Appl
43	532	55.5	123	4	US-10-837-217-17	Sequence 17, Appl
44	532	55.5	179	4	US-10-777-053-20	Sequence 20, Appl
45	532	55.5	179	4	US-10-837-217-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-849-602-30

; Sequence 30, Application US/09849602

; Publication No. US20030165834A1

; GENERAL INFORMATION:

; APPLICANT: Scania, Matthew J.

; APPLICANT: Old, Lloyd J.

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Chen, Yao-Tseng

; TITLE OF INVENTION: Colon Cancer Antigen Panel

; FILE REFERENCE: L0461/7105(JRV)

; CURRENT APPLICATION NUMBER: US/09/849,602

; CURRENT FILING DATE: 2001-05-04

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 30

; LENGTH: 180

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-849-602-30

Query Match	100.0%;	Score 959;	DB 3;	Length 180;
Best Local Similarity	100.0%;	Pred. No. 1.8e-67;		
Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MQAEGRTGGTGDADPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA	60	
Db	1	MQAEGRTGGTGDADPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA	60	
QY	61	PRGPHGGAASGLNGCCRGARGPESRLLFFYLAMPATPMEAEARRSLAQADAPLPVPG	120	
Db	61	PRGPHGGAASGLNGCCRGARGPESRLLFFYLAMPATPMEAEARRSLAQADAPLPVPG	120	
QY	121	VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCLQLSLMLWITQCFLPVFLAQPSPGQRR	180	
Db	121	VLLKEFTVSGNLTIRLTAAADHRQLQSLSSCLQLSLMLWITQCFLPVFLAQPSPGQRR	180	
RESULT 2				
US-10-207-655-71				
; Sequence 71, Application US/10207655				
; Publication No. US20030118592A1				
; GENERAL INFORMATION:				
; APPLICANT: Ledbetter, Jeffrey A.				
; APPLICANT: Hayden-Ledbetter, Martha S.				
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS				
; FILE REFERENCE: 390069.401C1				
; CURRENT APPLICATION NUMBER: US/10/207,655				
; CURRENT FILING DATE: 2002-07-25				
; NUMBER OF SEQ ID NOS: 426				


```
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-777-053-11

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||

RESULT 9
US-10-751-088-15
; Sequence 15, Application US/10751088
; Publication No. US20040158044A1
; GENERAL INFORMATION:
; APPLICANT: Knuth, Alexander; Jager, Elke; Chen, Yao, Scanlan, Matt;
; Gure, Ali, Old, Lloyd, Ritter, Gerd
; TITLE OF INVENTION: ISOLATED PEPTIDES CORRESPONDING TO AMINO ACID
; SEQUENCES OF NY-ESO-1, WHICH BIND TO MHC CLASS I AND MHC CL
; USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FULBRIGHT & JAWORSKI LLP
; STREET: 666 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10158
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10751,088
; FILING DATE: 02-Jan-2004
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/165,546D
; FILING DATE: 02-Oct-1998
; APPLICATION NUMBER: 09/062,422
; FILING DATE: April 17, 1998
; APPLICATION NUMBER: 08/937,263
; FILING DATE: September 15, 1997
; APPLICATION NUMBER: US 08/725,182
; FILING DATE: October 3, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, Norman D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 2166.4 CIP (09807811)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 318-3000
; TELEFAX: (212) 318-3400
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 180 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15
;
; JS-10-751-088-15
```

```
Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||

RESULT 10
US-10-657-022-74
; Sequence 74, Application US/10657022
; Publication No. US20040180354A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Liu, Liping
; APPLICANT: Liu, Zheng
; TITLE OF INVENTION: EPI TOPE SEQUENCES
; FILE REFERENCE: MANNK.032A
; CURRENT APPLICATION NUMBER: US/10/657,022
; CURRENT FILING DATE: 2003-09-04
; PRIOR APPLICATION NUMBER: 60/409123
; PRIOR FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 610
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-657-022-74

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
Db 1 MQAEGRTGGTGADGPGGPGIPDPGGNAGGPGGAGATGGRGPRGAGAAASGPGGGA 60
    |||
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPLPVP 120
    |||
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLLMWITQCFLPVLFAQPPSGQRR 180
    |||

RESULT 11
US-10-837-217-11
; Sequence 11, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPI TOPE OF
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
```

```
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-837-217-11

Query Match      100.0%; Score 959; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 12
US-10-877-373-9
; Sequence 9, Application US/10877373
; Publication No. US20040234541A1
; GENERAL INFORMATION:
; APPLICANT: Leth,, Bernard
; APPLICANT: Lucas, Sophie
; APPLICANT: De Smet, Charles
; APPLICANT: Godelaine, Daniele
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: LAGE-1 TUMOR ASSOCIATED NUCLEIC ACIDS
; FILE REFERENCE: L0461/7066
; CURRENT APPLICATION NUMBER: US/10/877,373
; CURRENT FILING DATE: 2004-06-25
; PRIOR APPLICATION NUMBER: US/09/341,829
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US 08/791,495
; PRIOR APPLICATION NUMBER: PCT/US98/01445
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-877-373-9

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 13
US-10-723-860-1270
; Sequence 1270, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1270
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1270

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 14
US-10-871-708-7
; Sequence 7, Application US/10871708
; Publication No. US20050118186A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Chih-Sheng
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; FILE REFERENCE: MANWK.035A
; CURRENT APPLICATION NUMBER: US/10/871,708
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: 60/479,554
; PRIOR FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Autoimmunogenic Cancer
US-10-871-708-7

Query Match      100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
Db 1 MQAEGRTGGTGDADGPGGPGIPDGGNAGGPGGAGATGGRPGAGAAARASGPGGGA 60
QY 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEARRSLAQDAPPLPVPG 120
QY 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKEFTVSGNLTIRLTAADHRQLQLSISSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
```

RESULT 15
US-10-895-523-3
; Sequence 3, Application US/10895523
; Publication No. US20050130920A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Lei, Xiang-Dong
; APPLICANT: Diamond, David C.
; TITLE OF INVENTION: EPITOPE SYNCHRONIZATION IN ANTIGEN
; TITLE OF INVENTION: PRESENTING CELLS
; FILE REFERENCE: MANNK.021CF1CC1
; CURRENT APPLICATION NUMBER: US/10/895,523
; CURRENT FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: 10/026,066
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 10/005,905
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/561,074
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/560,465
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,572
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,571
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: PCT/US01/13806
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 09/999,186
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-895-523-3

Query Match 100.0%; Score 959; DB 5; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.8e-67;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQAEGRTGGSTGDADGGGGIPDGGGNAGGPGAGATGGPRGAGAGARASGPGGGA 60
Db 1 MQAEGRTGGSTGDADGGGGIPDGGGNAGGPGAGATGGPRGAGAGARASGPGGGA 60

Qy 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPMPFATPMEAEIARRSLAQDAPPLPVP 120
Db 61 PRGPHGGAASGLNGCCRCGARGPESRLLEFYLPMPFATPMEAEIARRSLAQDAPPLPVP 120

Qy 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVLAPPSGQR 180
Db 121 VLLKEFTVSGNLTIRLTAAADHRQLQLSISSCLQQLSLLMWITQCFLPVLAPPSGQR 180

Search completed: February 4, 2006, 12:32:52
Job time : 175 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2006, 12:30:03 ; Search time 16 Seconds
(without alignments)
131.828 Million cell updates/sec

Title: US-09-529-206E-4

Perfect score: 959

Sequence: 1 MQAERGTGCGTGDADPGG.....WITQCFPVLAQPPSGQRR 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88029 seqs, 11718060 residues

Total number of hits satisfying chosen parameters: 88029

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA.New.*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pap.*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pap.*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap.*
4: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pap.*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pap.*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pap.*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pap.*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	959	100.0	180	7	US-11-155-288-7
2	954	99.5	240	7	US-11-021-441-28
3	155.5	16.2	641	6	US-10-848-976-1
4	153	16.0	1079	7	US-11-052-554A-145
5	151.5	15.8	1464	7	US-11-000-463-243
6	151.5	15.8	1464	7	US-11-186-284-28
7	151.5	15.8	1464	7	US-11-021-603-2
8	151.5	15.8	1467	6	US-10-821-234-1096
9	151	15.7	1901	7	US-11-052-554A-135
10	149.5	15.6	615	7	US-11-052-554A-132
11	147	15.3	1381	7	US-11-052-554A-138
12	146.5	15.3	1767	6	US-10-995-561-911
13	146.5	15.3	1767	6	US-10-995-561-914
14	146.5	15.3	1806	6	US-10-995-561-912
15	146.5	15.3	1806	6	US-10-995-561-915
16	146.5	15.3	1818	6	US-10-995-561-910
17	146.5	15.3	1818	6	US-10-995-561-913
18	146	15.2	853	7	US-11-052-554A-149
19	145.5	15.2	1306	7	US-11-052-554A-139
20	145	15.1	741	7	US-11-052-554A-161
21	145	15.1	1329	7	US-11-052-554A-136
22	143	14.9	1660	7	US-11-052-554A-137
23	142.5	14.9	484	7	US-11-052-554A-170
24	142.5	14.9	1466	7	US-11-186-284-33
25	140.5	14.7	606	7	US-11-052-554A-163

26 138 14.4 1538 7 US-11-052-554A-146 Sequence 146, App
27 137.5 14.3 801 7 US-11-052-554A-166 Sequence 166, App
28 137.5 14.3 1166 6 US-10-821-234-964 Sequence 964, App
29 135.5 14.1 749 7 US-11-052-554A-148 Sequence 148, App
30 135.5 14.1 1366 6 US-10-821-234-1431 Sequence 1431, App
31 135.5 14.1 1366 7 US-11-186-284-31 Sequence 31, App
32 134.5 14.0 914 7 US-11-052-554A-160 Sequence 160, App
33 134 14.0 80 7 US-11-029-003-34 Sequence 34, App
34 133.5 13.9 3063 7 US-11-186-284-26 Sequence 26, App
35 133 13.9 767 7 US-11-052-554A-154 Sequence 154, App
36 131.5 13.7 1736 7 US-11-124-368A-329 Sequence 329, App
37 131 13.7 200 6 US-10-841-129-7 Sequence 7, App
38 131 13.7 200 7 US-11-060-659-32 Sequence 32, App
39 130.5 13.6 434 7 US-11-052-554A-167 Sequence 167, App
40 130.5 13.6 576 7 US-11-052-554A-164 Sequence 164, App
41 130.5 13.6 923 7 US-11-052-554A-147 Sequence 147, App
42 129 13.5 639 7 US-11-052-554A-165 Sequence 165, App
43 129 13.5 1496 7 US-11-186-284-35 Sequence 35, App
44 128.5 13.4 778 7 US-11-052-554A-144 Sequence 144, App
45 128 13.3 618 7 US-11-052-554A-150 Sequence 150, App

ALIGNMENTS

RESULT 1

US-11-155-288-7
; Sequence 7, Application US/11155288
; Publication No. US20060008468A1
; GENERAL INFORMATION:
; APPLICANT: Chiang, Chih-Sheng
; APPLICANT: Simard, John J. L.
; TITLE OF INVENTION: COMBINATIONS OF TUMOR-ASSOCIATED
; TITLE OF INVENTION: ANTIGENS IN DIAGNOSTICS FOR VARIOUS TYPES OF CANCERS
; FILE REFERENCE: MANWK.050A
; CURRENT APPLICATION NUMBER: US/11/155,288
; CURRENT FILING DATE: 2005-06-17
; PRIOR APPLICATION NUMBER: 60/580,969
; PRIOR FILING DATE: 2004-06-17
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-155-288-7

Query Match 100.0%; Score 959; DB 7; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.3e-78;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MQAERGTGCGTGDADPGGCGIPDGPAGNAGGCGEAGATGGGPKGAGARASGPGGGA 60
Db 1 MQAERGTGCGTGDADPGGCGIPDGPAGNAGGCGEAGATGGGPKGAGARASGPGGGA 60
Qy 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEALARRSLAQADAPPLVPVG 120
Db 61 PRGPHGAASGLNGCCRCGARGPESRLLEFYLLAMPATPMEALARRSLAQADAPPLVPVG 120
Qy 121 VLLKFTVSGNLTIRLTAAHHRQLSLSSCLQLSLMWITQCFLPVFLAQPSPGQRR 180
Db 121 VLLKFTVSGNLTIRLTAAHHRQLSLSSCLQLSLMWITQCFLPVFLAQPSPGQRR 180

RESULT 2

US-11-021-441-28
; Sequence 28, Application US/11021441
; Publication No. US20050249748A1
; GENERAL INFORMATION:
; APPLICANT: DUBENSKI, Thomas W., Jr.
; APPLICANT: FORTNOY, Daniel A.
; APPLICANT: LUCKETT, William S., Jr.
; APPLICANT: COOK, David N.


```

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 243
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-243

Query Match      15.8%; Score 151.5; DB 7; Length 1464;
Best Local Similarity 37.2%; Pred. No. 5.5e-06;
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;

QY      3 AEGR-GTGGSTGDA--DGGPGGIPDPGPGNAGGPGCEAGATGGRGPRGA-----GAARASG 55
Db      825 AKGEPDAGAKGDAGPPGPGAGPAGPGPTGNVGAPGAKGARGSGAGPPGATGFFGAAAGRVG 884

QY      56 P-----GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQ 111
Db      885 PPGPSGNAGPPGPPGAG-----KEGKGPRGE-----TCPAGRPGEV----- 922

QY      112 DAPPLPVP 120
Db      923 -GPPGP-PG 929

RESULT 6
US-11-186-284-28
; Sequence 28, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burtgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY FOR COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-28

Query Match      15.8%; Score 151.5; DB 7; Length 1464;
Best Local Similarity 37.2%; Pred. No. 5.5e-06;
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;

QY      3 AEGR-GTGGSTGDA--DGGPGGIPDPGPGNAGGPGCEAGATGGRGPRGA-----GAARASG 55
Db      825 AKGEPDAGAKGDAGPPGPGAGPAGPGPTGNVGAPGAKGARGSGAGPPGATGFFGAAAGRVG 884

QY      56 P-----GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQ 111
Db      885 PPGPSGNAGPPGPPGAG-----KEGKGPRGE-----TCPAGRPGEV----- 922

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 243
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-000-463-243

Query Match      15.8%; Score 151.5; DB 7; Length 1464;
Best Local Similarity 37.2%; Pred. No. 5.5e-06;
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;

QY      3 AEGR-GTGGSTGDA--DGGPGGIPDPGPGNAGGPGCEAGATGGRGPRGA-----GAARASG 55
Db      825 AKGEPDAGAKGDAGPPGPGAGPAGPGPTGNVGAPGAKGARGSGAGPPGATGFFGAAAGRVG 884

QY      56 P-----GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQ 111
Db      885 PPGPSGNAGPPGPPGAG-----KEGKGPRGE-----TCPAGRPGEV----- 922

; Sequence 2, Application US/11021603
; Publication No. US20060003954A1
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajinder
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 06275-254US1
; CURRENT APPLICATION NUMBER: US/11/021,603
; CURRENT FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: US/10/149,352
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: PCT/GB00/04741
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: GB 9929487.8
; PRIOR FILING DATE: 1999-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 4.0
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-021-603-2

Query Match      15.8%; Score 151.5; DB 7; Length 1464;
Best Local Similarity 37.2%; Pred. No. 5.5e-06;
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;

QY      3 AEGR-GTGGSTGDA--DGGPGGIPDPGPGNAGGPGCEAGATGGRGPRGA-----GAARASG 55
Db      825 AKGEPDAGAKGDAGPPGPGAGPAGPGPTGNVGAPGAKGARGSGAGPPGATGFFGAAAGRVG 884

QY      56 P-----GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFYLAMPPATPMEAEIARRSLAQ 111
Db      885 PPGPSGNAGPPGPPGAG-----KEGKGPRGE-----TCPAGRPGEV----- 922

; Sequence 1096, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 1096
; LENGTH: 1467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1096

Query Match      15.8%; Score 151.5; DB 6; Length 1467;
Best Local Similarity 37.2%; Pred. No. 5.5e-06;
```

```
Matches 48; Conservative 4; Mismatches 42; Indels 35; Gaps 8;
QY 3 AGR-GTGSGTDA--DGGCGGCGIIPDGGNAGGPGGAGATGGRPRGA-----GAARAG 55
Db 828 AKGEPDAGAKGDAGPPGAGPAGPPGPIGNVGAPGAKGARGAGPPGATGFPGAAGRVG 887
QY 56 P-----GGGAPRPHGGAASGLNGCCRCGARGPESRLLEFFLAMPFPATPMEALARRSLAQ 111
Db 888 PPGPSNAGPPGPPGAG-----KGGKGRGE-----TGPAGRPGEV----- 925
QY 112 DAPPLEVPG 120
Db 926 -GPPGP-PG 932

RESULT 9
US-11-052-554A-135
; Sequence 135, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 135
; TYPE: PR
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-135

Query Match 15.7%; Score 151; DB 7; Length 1901;
Best Local Similarity 41.1%; Pred. No. 8e-06;
Matches 37; Conservative 2; Mismatches 31; Indels 20; Gaps 3;
QY 5 GRGTGGSTGDADGPGGP-----GIPDGGNAGGPGGEEA-----GATGGRPRGA 48
Db 140 GAGAGAPGAGGPGGPGAGLWNGGPGGAGSGGTGGAGGAGGWLFGVGGAGVGGAGG 199
QY 49 GAARAGSPG-----GGAPRPHGGAASGLNG 74
Db 200 GTGAGGPGGLWGGGAGGAGVGGAGGTTGG 229

RESULT 10
US-11-052-554A-152
; Sequence 152, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 152
; TYPE: PR
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-152
```

```
Query Match 15.6%; Score 149.5; DB 7; Length 615;
Best Local Similarity 49.3%; Pred. No. 3.3e-06;
Matches 35; Conservative 5; Mismatches 26; Indels 5; Gaps 3;
QY 7 GTGGSTGDADGPGGPIPDGGNAGG--PGBAGATGGRGRGA-GAARASPGGGAPRG 63
Db 532 GTGGNGDGGAGGAPGL--GGAGGAGGWLIGQSGTGGGAGGAGGAGGAGGSGGAG 589
QY 64 PHGGAASGLNG 74
Db 590 GHGDTTSGKNG 600

RESULT 11
US-11-052-554A-138
; Sequence 138, Application US/11052554A
; Publication No. US20050288866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 138
; LENGTH: 1381
; TYPE: PR
; ORGANISM: Mycobacterium tuberculosis H37Rv
US-11-052-554A-138

Query Match 15.3%; Score 147; DB 7; Length 1381;
Best Local Similarity 36.1%; Pred. No. 1.3e-05;
Matches 39; Conservative 2; Mismatches 39; Indels 28; Gaps 4;
QY 5 GRGTGGSTG-----DADGPGGPIP-----DPPGNAGGPGGAGATGGRPRGAAGA 53
Db 523 GNGTGVNGADNTLNPDTFCGAGGPGGAGGAGGAGGAGGPGGTGTTGGNGGNGGNG 582
QY 54 SFPGGG-----APRPHGGA-----ASGLNGCCRCGARGPE 84
Db 583 GNGGNGGNGGNAGNNSTNAPVGGGAGGAGGAGGAGGAANGGTAGSQ 630

RESULT 12
US-10-995-561-911
; Sequence 911, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 911
; LENGTH: 1767
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-995-561-911

Query Match 15.3%; Score 146.5; DB 6; Length 1767;
Best Local Similarity 35.4%; Pred. No. 1.9e-05;
Matches 45; Conservative 8; Mismatches 45; Indels 29; Gaps 6;
```



```
QY 7 GTGGSTGDADGPGGPGI-----PDGPGNAGGPGGAGATGGRGPRGA-GAARASG-P 56
Db 1291 GVGGDKGEDGDPGQPPGPPGSGAGPPGPKRGPPGAAGAGRQGEKGAKEAGAGGPP 1350
QY 57 GGGAPRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPL 116
Db 1351 GKTGVPVPGQGA-----GKPGPEG-----LRGIPGVGQGLPGAAGQDGPFG 1393
QY 117 PV--PGV 121
Db 1394 PMGPPGL 1400
```

```
RESULT 13
US-10-995-561-914
; Sequence 914, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 914
; LENGTH: 1767
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-914
```

```
Query Match 15.3%; Score 146.5; DB 6; Length 1767;
Best Local Similarity 35.4%; Pred. No. 1.9e-05;
Matches 45; Conservative 8; Mismatches 45; Indels 29; Gaps 6;

QY 7 GTGGSTGDADGPGGPGI-----PDGPGNAGGPGGAGATGGRGPRGA-GAARASG-P 56
Db 1291 GVGGDKGEDGDPGQPPGPPGSGAGPPGPKRGPPGAAGAGRQGEKGAKEAGAGGPP 1350
QY 57 GGGAPRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPL 116
Db 1351 GKTGVPVPGQGA-----GKPGPEG-----LRGIPGVGQGLPGAAGQDGPFG 1393
QY 117 PV--PGV 121
Db 1394 PMGPPGL 1400
```

```
RESULT 14
US-10-995-561-912
; Sequence 912, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 912
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-912
```

```
Query Match 15.3%; Score 146.5; DB 6; Length 1806;
Best Local Similarity 35.4%; Pred. No. 1.9e-05;
```

```
Matches 45; Conservative 8; Mismatches 45; Indels 29; Gaps 6;

QY 7 GTGGSTGDADGPGGPGI-----PDGPGNAGGPGGAGATGGRGPRGA-GAARASG-P 56
Db 1330 GVGGDKGEDGDPGQPPGPPGSGAGPPGPKRGPPGAAGAGRQGEKGAKEAGAGGPP 1389
QY 57 GGGAPRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPL 116
Db 1390 GKTGVPVPGQGA-----GKPGPEG-----LRGIPGVGQGLPGAAGQDGPFG 1432
QY 117 PV--PGV 121
Db 1433 PMGPPGL 1439
```

```
RESULT 15
US-10-995-561-915
; Sequence 915, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 915
; LENGTH: 1806
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-915
```

```
Query Match 15.3%; Score 146.5; DB 6; Length 1806;
Best Local Similarity 35.4%; Pred. No. 1.9e-05;
Matches 45; Conservative 8; Mismatches 45; Indels 29; Gaps 6;

QY 7 GTGGSTGDADGPGGPGI-----PDGPGNAGGPGGAGATGGRGPRGA-GAARASG-P 56
Db 1330 GVGGDKGEDGDPGQPPGPPGSGAGPPGPKRGPPGAAGAGRQGEKGAKEAGAGGPP 1389
QY 57 GGGAPRGPHGGAASGLNGCCRCGARGPESRLLEFYLLAMPFATPMEAEALARRSLAQDAPPL 116
Db 1390 GKTGVPVPGQGA-----GKPGPEG-----LRGIPGVGQGLPGAAGQDGPFG 1432
QY 117 PV--PGV 121
Db 1433 PMGPPGL 1439
```

```
Search completed: February 4, 2006, 12:33:14
Job time : 17 secs
```

This Page Blank (uspto)